



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 273 835**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01929654 .0**

86 Fecha de presentación : **27.04.2001**

87 Número de publicación de la solicitud: **1284296**

87 Fecha de publicación de la solicitud: **19.02.2003**

54

Título: **Método para la detección de genomas minoritarios en cuasiespecies virales usando microchips de ADN.**

30

Prioridad: **27.04.2000 ES 200001068**

45

Fecha de publicación de la mención BOPI:
16.05.2007

45

Fecha de la publicación del folleto de la patente:
16.05.2007

73

Titular/es:
**Consejo Superior de Investigaciones Científicas
c/ Serrano, 117
28006 Madrid, ES**

72

Inventor/es: **Arias Esteban, Armando;
Baranowski, Eric;
Briones Llorente, Carlos;
Domingo Solans, Esteban;
Escarmis Homs, Cristina;
Gómez Castilla, Jordi;
Martín Ruiz-Jarabo, Carmen y
Parro García, Víctor**

74

Agente: **Arias Sanz, Juan**

ES 2 273 835 T3

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

DESCRIPCIÓN

Método para la detección de genomas minoritarios en cuasiespecies virales usando microchips de ADN.

5 **Sector de la técnica**

La invención se refiere a un método para la detección de genomas minoritarios en cuasiespecies virales. La técnica debería usarse para el diagnóstico genético de enfermedades virales mediante el análisis de cuasiespecies de virus patógenos empleando microchips de ADN u otras técnicas.

10 **Estado de la técnica**

Los virus ARN (que contienen ácido ribonucleico o ARN como material genético) se han asociado a numerosas enfermedades que afectan a humanos, animales y plantas y que causan elevada mortalidad y tienen cuantiosas pérdidas económicas. La asociación de virus ARN con enfermedades importantes es bien conocida (véanse, los capítulos y referencias de los siguientes libros: Mohanty *et al.*, 1981; Robart, 1995; Fields *et al.*, 1996; Granoff *et al.*, 1999; Flint *et al.*, 2000). Se sabe que los virus ARN son responsables tanto de infecciones agudas (con síntomas de una duración de varias horas o días) como de infecciones crónicas o persistentes (con o sin síntomas aparentes, pero con virus detectable durante semanas, meses, años o décadas). Ejemplos de infecciones humanas agudas debidas a virus ARN son la gripe, el sarampión, el resfriado común, la poliomieltis, varios tipos de encefalitis víricas y de fiebres hemorrágicas y las asociadas con el virus de la hepatitis tipo A (virus de la familia *Picornaviridae*). Ejemplos de infecciones animales agudas debidas a virus ARN son la fiebre aftosa, la estomatitis vesicular, enteritis víricas, enfermedad de Teschen, encefalomielitis aviar, entre otras. Ejemplos de enfermedades humanas crónicas debidas a virus ARN son hepatitis tipo C y SIDA (asociada a los virus de la inmunodeficiencia humana). Dentro de este grupo pueden incluirse síndromes asociados a la persistencia de ciertos virus tras una infección aguda; ejemplos de éstos son el síndrome post-polio (resultado de la persistencia de formas genómicas variantes del virus de la poliomieltis) y la panencefalitis esclerosante subaguda (resultado de la persistencia de formas genómicas variantes del virus del sarampión). Ejemplos de infecciones animales crónicas debidas a virus ARN son meningitis causadas por el virus de la cariolinfomeningitis infecciosa o anemia equina causada por el virus de la anemia infecciosa equina.

El material genético de los virus ARN es ARN (esto es, el material que se multiplica en el interior de las células infectadas y que es responsable de las propiedades biológicas y de virulencia de un virus). Éste contiene información o un mensaje genético en forma de un polinucleótido resultado de la polimerización de cuatro nucleótidos [adenosina monofosfato (AMP), citosina monofosfato (CMP), guanosina monofosfato (GMP) y uridina monofosfato (UMP)]. Polímeros de estos cuatro monómeros, de longitud comprendida entre 3.000 y 4.000 unidades, constituyen el material genético de todos los virus ARN descritos hasta la fecha. Durante los procesos de infección (multiplicación de un virus en células u organismos) las poblaciones virales frecuentemente constituyen colectivos con un elevado número de partículas infecciosas que pueden alcanzar desde 10^5 hasta 10^{12} partículas infecciosas en el momento del pico de una infección aguda o en varios puntos durante una infección crónica (Domingo *et al.*, 1999d). El mensaje genético (representado en la secuencia de nucleótidos del ARN genómico viral) no es idéntico en los genomas individuales que componen la población vírica, sino que muchos genomas individuales difieren del resto en una o más posiciones. Estos comprenden un conjunto de secuencias estrechamente relacionadas pero no idénticas y a este conjunto de secuencias se le denomina cuasiespecie vírica. El concepto de cuasiespecies aplicado a virus ARN tiene una base teórica y una base experimental.

El concepto teórico de cuasiespecies arranca de un trabajo teórico de M. Eigen concerniente a un sistema de moléculas replicantes con una producción continua de errores debido a la fidelidad limitada en el reconocimiento de los nucleótidos del molde (Eigen, 1971). En este primer trabajo se distinguió el concepto de “auto-instrucción” del concepto general de autocatálisis; el concepto de “auto-instrucción” se propuso como necesario para la actividad de una molécula como molde para la replicación. Según ésta, y en el caso de ARN, los nucleótidos del ARN molde dictan la incorporación de sus respectivos complementarios durante el proceso de replicación. Por tanto, adenina (A) es complementaria a uracilo (U) y guanina (G) es complementaria a citosina (C). Se definió un factor de calidad que representa la fracción de los procesos de copia que dan lugar a una copia exacta al molde. Cuando la fidelidad de copia no es igual a 1 (el valor máximo posible) la copia de la secuencia principal, también denominada secuencia maestra, producirá moléculas erróneas que tendrán una cierta probabilidad de distribución. Por tanto, dependiendo de la fidelidad de copia habrá distinta abundancia de genomas con una, dos, tres o más diferencias respecto a la secuencia maestra.

En este primer trabajo, M. Eigen se refiere a la “cola de cometa” de copias erróneas, mientras que los términos “cuasiespecies” y “espectro mutante” aparecieron posteriormente en la literatura con el trabajo de M. Eigen y P. Schuster (Eigen *et al.*, 1977; Eigen *et al.*, 1978a; Eigen *et al.*, 1978b; Eigen *et al.*, 1979). En estos trabajos, se desarrolló el concepto de equilibrio de selección de las múltiples variantes generadas en un proceso de copia con fidelidad limitada. Este equilibrio es generalmente metaestable, en el sentido de que puede colapsar cuando un mutante ventajoso aparece en la población. Este colapso da lugar a una reorganización de las variantes de la población y un nuevo punto de equilibrio de selección. Como se justificará más adelante, cambios y fluctuaciones en el equilibrio de los variantes genómicos que constituyen la población ocurren frecuentemente en un virus ARN y son relevantes para la invención que se presenta aquí. El desarrollo teórico del concepto de cuasiespecies ha sido descrito en numerosos trabajos y

revisiones posteriores (Eigen *et al.*, 1988; Eigen, 1992; Eigen, 1996; Schuster *et al.*, 1999; Domingo *et al.*, 2000). El modelo teórico de cuasiespecies representa la combinación de los principios de evolución darwiniana y de la teoría de la información, y resultó fundamental para entender la dinámica poblacional de los virus ARN, como se describe a continuación.

5

La primera evidencia experimental de que los virus ARN mostraban propiedades típicas de cuasiespecies fue obtenida en trabajos con el bacteriófago (virus bacteriano) Q β que infecta ciertas cepas de la bacteria *Escherichia coli* (Domingo *et al.*, 1976; Batschelet *et al.*, 1976; Domingo *et al.*, 1978). Las observaciones más relevantes fueron reveladas por el estudio de la tasa de reversión (conversión a la secuencia genómica inicial, también denominada de “tipo salvaje”) de un mutante del bacteriófago Q β durante su multiplicación en *Escherichia coli*. Esta tasa de reversión fue del orden de 10⁻⁴ sustituciones por nucleótido copiado; esto es, de cada 10.000 veces que la enzima encargada de copiar el ARN mutante de Q β pasaba por la posición mutada del molde, el producto tenía una base distinta de la complementaria que se supone le corresponde. Los experimentos y los cálculos que conducen a esta conclusión han sido descritos por Domingo *et al.*, y Batschelet *et al.* (Domingo *et al.*, 1976; Batschelet *et al.*, 1976). La segunda observación relevante fue el descubrimiento de una elevada heterogeneidad genética (presencia de distribuciones de mutantes diferentes en una o más posiciones) en las poblaciones del bacteriófago Q β (Domingo *et al.*, 1978). Los experimentos decisivos consistieron en aislar bacteriófago Q β procedente de una placa individual y aislada formada sobre la superficie de una capa de *Escherichia coli* susceptible (se denomina “placa” a la zona de células muertas producidas por el virus que se multiplica a partir de una partícula infecciosa única inicial).

20

Experimentos clásicos de virología empleando diluciones de material infectado han demostrado que en el bacteriófago Q β y otros muchos virus ARN, una placa se origina a partir de un solo genoma ARN infeccioso y no de varios (una recopilación de estas pruebas aparece en Luria *et al.*, 1978). Cuando un virus proveniente de una placa se propaga en la bacteria *Escherichia coli*, se formaron poblaciones que, al ser analizadas por métodos de clonaje biológico (aislamiento de placas virales individuales bien separadas) resultaron ser genéticamente heterogéneas. El procedimiento consistió en marcar virus de ARN obtenido de placas individuales con ³²P-fosfato (fosfato neutro, alguno de cuyos átomos de fósforo normales habían sido sustituidos por la forma radioactiva o isótopo-32) y analizar el ARN tras hidrólisis con ribonucleasa T1 (una enzima que rompe ARNs de cadena sencilla en las posiciones ocupadas por ácido guanílico GMP). El análisis consiste en una electroforesis bidimensional (resultado de realizar dos electroforesis sucesivas con campos eléctricos perpendiculares) en la que las posiciones de las manchas de radioactividad correspondientes a los oligonucleótidos T1 (los productos de digestión del ARN viral con ribonucleasa T1) dependen de su composición de nucleótidos exacta. Por ello, la posición de los oligonucleótidos tras electroforesis bidimensional permite detectar e identificar mutaciones (De Watcher *et al.*, 1972). El análisis de ARN de clones individuales que se habían originado a partir de una sola secuencia de ARN demostró que la mayoría de ellos diferían en una o más posiciones con respecto al genoma inicial (parental) de la población, y que sólo el 14% de los genomas eran idénticos a los genomas originales (Domingo *et al.*, 1978). Según estos resultados, el bacteriófago Q β replicaba con una elevada tasa de producción de copias erróneas y que sus poblaciones, en vez de ser homogéneas, contenían distribuciones dinámicas de variantes.

40

Tal como se describió en el artículo original (Domingo *et al.*, 1978): “Una población del fago Q β se halla en equilibrio dinámico con mutantes viables, que, por una parte aparecen con alta frecuencia, y por otra se seleccionan negativamente. El genoma no puede describirse como una secuencia única sino como un promedio ponderado de un gran número de secuencias individuales”.

45

Trabajos posteriores empleando numerosos virus ARN tanto de animales como de plantas, han demostrado que todos los virus analizados tienen las mismas características de estructura poblacional que aquellas descritas arriba para el bacteriófago Q β . En las últimas dos décadas, se han mejorado los métodos de clonaje de genomas víricos (biológicos y moleculares) y también los procedimientos de secuenciación rápida de nucleótidos, incluida la secuenciación automatizada (como manuales explicativos de las nuevas tecnologías de recombinación de ADN *in vitro* y secuenciación, véase Sambrook *et al.*, 1989; Howe *et al.*, 1989; Heitman, 1993).

50

La aplicación de estas técnicas al análisis molecular de genomas de ARN ha resultado en la determinación de la estructura de cuasiespecies de los virus ARN que infectan humanos, animales y plantas. Dos tipos de resultados han reafirmado la naturaleza en cuasiespecies de virus ARN y de otros elementos genéticos en los que el ARN interviene en su ciclo replicativo: el cálculo de altas tasas de mutación por técnicas tanto genéticas como bioquímicas (tasas generalmente comprendidas entre 10⁻³ y 10⁻⁵ sustituciones por nucleótido copiado) y la demostración directa de la presencia de un espectro de mutantes en las poblaciones virales.

55

Además de los virus con genoma de ARN, se ha demostrado que también poseen una elevada tasa de mutación y estructura de cuasiespecie aquellos virus con genoma de ADN en cuyo ciclo infeccioso se genera ARN viral como uno de los intermedios replicativos. Algunos de los virus con genoma de ADN y estructura de cuasiespecie que infectan a humanos y animales pertenecen a la familia *Hepadnaviridae*. El más conocido y estudiado de éstos es el virus de la hepatitis B humana (VHB) (las características generales de los hepadnavirus se pueden consultar en Fields *et al.* 1996). Se ha estimado que entre un 5 y un 10% de la población mundial es portadora del VHB, existiendo zonas geográficas donde su incidencia se considera endémica, como África o el sudeste asiático (Maynard, 1990; Coleman *et al.*, 1998). Entre el 5 y el 10% de los individuos adultos que se exponen al VHB se convierten en portadores crónicos, pudiendo desarrollar cirrosis y cáncer de hígado, que es fatal en aproximadamente la mitad de ellos (Liaw *et al.*, 1988).

65

Existen considerables pruebas experimentales para las altas tasas de mutación y la heterogeneidad en poblaciones de ARN; véase, por ejemplo: Borrow *et al.*, 1997; Borrow *et al.*, 1998; Briones *et al.*, 2000; Chen *et al.*, 1996; Cornelissen *et al.*, 1997; Domingo *et al.*, 1993; Domingo, 1996; Domingo, 1997b; Domingo *et al.*, 1999a; Domingo, 1999d; Domingo *et al.*, 2000; Eigen, 1996; Escarmís *et al.*, 1999; Escarmís *et al.*, 1996; Flint *et al.*, 2000; Granoff *et al.*, 1999; Mateu *et al.*, 1989; Morse, 1993; Morse, 1994; Mortara *et al.*, 1998; Nájera *et al.*, 1995; Quiñones-Mateu *et al.*, 1996a; Quiñones-Mateu *et al.*, 1996b; Ruiz-Jarabo *et al.*, 1999; Taboga *et al.*, 1997; Weidt *et al.*, 1995; y Weiner *et al.*, 1995.

A partir de estos trabajos se puede concluir que las poblaciones de virus con genomas de ARN y aquellas que utilizan el ARN como una molécula intermedia dentro de su ciclo replicativo tienen un comportamiento de cuasiespecies y que dicho comportamiento es importante para la adaptabilidad, supervivencia y poder patogénico de los virus. En la última edición de la Enciclopedia de Virología del año 1999 se da la siguiente definición generalizada de cuasiespecies, actualmente empleada en virología: “Cuasiespecies son distribuciones dinámicas de genomas mutantes recombinantes que no son idénticos pero que están estrechamente relacionados y que están sujetos a continuos procesos de variación genética, competición y selección, y que actúan como una unidad de selección” (Domingo, 1999a).

La estructura en cuasiespecies de los virus ARN (y los que utilizan el ARN como intermedio replicativo) tiene numerosas implicaciones biológicas que han sido revisadas en numerosos libros y monografías (véase, por ejemplo, Morse, 1993; Morse, 1994; Gibbs *et al.*, 1995; Domingo *et al.*, 1999d; Domingo *et al.*, 2000). Algunas de las implicaciones biológicas son relevantes para esta invención y se indican a continuación:

1) Las cuasiespecies víricas constituyen reservas de variantes genéticos (y fenotípicos; esto es, una variante de comportamiento biológico) que tienen una cierta probabilidad de ser seleccionados en respuesta a una selección aplicada desde el exterior del organismo o endógena del organismo infectado (Domingo, 1996; Forns *et al.*, 1999).

2) Entre los variantes que constituyen el espectro de mutantes de una cuasiespecie vírica se hallan mutantes con una sensibilidad disminuida a inhibidores empleados en la terapia de enfermedades víricas (Cornelissen *et al.*, 1997; Nájera *et al.*, 1995; Lech *et al.*, 1996; Quiñones-Mateu *et al.*, 1998; Havlir *et al.*, 1996). La presencia de variantes minoritarios con mutaciones que confieren distintos grados de resistencia a inhibidores es una de las causas que contribuyen al fracaso terapéutico en el tratamiento de infecciones por el virus de la inmunodeficiencia humana (entre los estudios que prueban este hecho están, Richman, 1994; Domingo *et al.*, 1997b; Palmer *et al.*, 1999). Para el virus de la inmunodeficiencia humana existen catálogos de mutaciones que, aisladas o en combinación, contribuyen a la baja eficacia de los tratamientos antirretrovirales (Schinazi *et al.*, 1997; Schinazi *et al.*, 1999; Menéndez-Arias *et al.*, 1999).

3) Entre los variantes que constituyen el espectro de mutantes de una cuasiespecie vírica se hallan mutantes con sensibilidad disminuida a anticuerpos o a células T citotóxicas (CTLs) (véase, por ejemplo, Borrow *et al.*, 1997; Borrow *et al.*, 1998; McMichael *et al.*, 1997; Weidt *et al.*, 1995; Weiner *et al.*, 1995; Domingo *et al.*, 1993; Taboga *et al.*, 1997; Mortara *et al.*, 1998).

4) En algunos casos se han obtenido pruebas directas de que la presencia de variantes antigénicas, o de otros tipos de variantes con propiedades biológicas alteradas, influyen en la progresión de la enfermedad viral in vivo (Pawlotsky *et al.*, 1998; Forns *et al.*, 1999; Evans *et al.*, 1999).

Como una consecuencia de la estructura de cuasiespecies de ciertos virus y de su rápida diversificación en la naturaleza, la mayoría de virus patógenos circulan como genomas distintos que se han dividido en tipos, subtipos, genotipos o biotipos y que pueden requerir reactivos específicos de diagnóstico. Estas subdivisiones de un virus se dan en importantes virus patógenos como es el virus de la inmunodeficiencia humana, virus de la hepatitis C, virus de la gripe, rotavirus humanos y animales, virus de la poliomielitis, virus de la fiebre aftosa o glosopeda y otros muchos (véase, por ejemplo, Murphy, 1996, y la solicitud de patente europea EP 0 984 067 A2).

Los métodos actuales para el diagnóstico molecular de virus se basan en la detección de genomas virales mayoritarios presentes en la población, empleando secuenciación directa de nucleótidos, métodos indirectos de detección de secuencias (por hibridación de ácidos nucleicos, polimorfismos revelados mediante empleo de enzimas de restricción sobre copias en ADN del ARN vírico, o por variaciones en la migración electroforética de heteroduplex formados por hibridación de un ADN de referencia con el ADN copia del genoma a analizar, etc).

La secuenciación de clones biológicos o moleculares proporciona una adecuada descripción genética de la cuasiespecie, aunque el muestreo de genomas analizables para cada cuasiespecie viral es limitado, generalmente no más de 20-30 clones (Briones *et al.*, 2000). Por otra parte, no todas las técnicas citadas resultan sensibles para la detección de genomas minoritarios dentro de la cuasiespecie. Por ejemplo, mediante el análisis de secuencias consenso se puede localizar un genoma minoritario siempre que esté presente en una proporción mayor o igual al 30-50%. Determinadas técnicas de hibridación de ácidos nucleicos permiten detectar genomas minoritarios cuando éstos están presentes en proporciones mayores del 10-20% y por secuenciación de 20 clones moleculares derivados de la cuasiespecie (técnica muy lenta y laboriosa) se puede llegar a detectar genomas presentes en una proporción mayor o igual al 5%.

El desarrollo reciente de la tecnología de microarrays de ADN, también llamados chips o microchips de ADN (Southern *et al.*, 1994; para una revisión véase Nature Genetics 21, suplemento, 1999), en la cual miles de sondas mo-

leculares, principalmente oligonucleótidos, se pueden fijar covalentemente a un soporte sólido (vidrio, nitrocelulosa, nylon etc.), ha permitido la identificación de polimorfismos de un solo nucleótido en tan solo una ronda de hibridación. Además, esta técnica posee mucha mayor sensibilidad que las citadas en el párrafo anterior, permitiendo detectar genomas minoritarios presentes en una proporción de tan sólo un 1% (Gerry *et al.*, 1999).

5

La tecnología de los chips de ADN explota la técnica desarrollada por E. Southern (Southern, 1975), según la cual los ácidos nucleicos se pueden fijar a un soporte sólido y formar híbridos estables con sus complementarios marcados radiactiva o fluorescentemente. La estabilidad de los híbridos depende del grado de complementariedad de la secuencia de nucleótidos y de factores externos tales como la fuerza iónica del medio, el pH o la temperatura. Es posible diseñar y sintetizar oligonucleótidos, por ejemplo de 10 a 30 nucleótidos, en los que alguna de sus posiciones, generalmente la posición central, difiere de la que está presente en la cadena complementaria de un gen determinado. Cuando las secuencias de dicho gen se marcan con un compuesto radiactivo o fluorescente y se ponen en contacto con un conjunto de oligonucleótidos idénticos de cuatro en cuatro excepto en la posición central, donde puede haber A, C, G o T, y en condiciones específicas de fuerza iónica, pH y temperatura, sólo se formarán híbridos estables sólo donde el apareamiento entre bases complementarias sea completo. Un resultado positivo en la hibridación identifica inmediatamente al nucleótido presente en la posición del gen bajo estudio (Hacia *et al.*, 1998).

10

Para la construcción de un microchip de ADN, se pueden seguir una de dos estrategias básicas: una de ellas consiste en depositar directamente una sonda previamente sintetizada sobre un soporte sólido. La sonda puede ser un oligonucleótido, un fragmento amplificado por PCR, un plásmido o un fragmento de ADN purificado. Otra estrategia consiste en sintetizar las sondas *in situ*, bien por el método de desprotección fotoquímica de nucleótidos inactivados por una sustancia fotolábil (patente norteamericana número 6.022.963), bien por el método de chorro de tinta (Blanchard *et al.*, 1996) o bien por confinamiento físico de los reactivos (Maskos y Southern, 1993). Para la deposición directa de la muestra se requiere un sistema robotizado ("arrayer") que permite imprimir hasta 2.500 muestras (de 100 micrometros de diámetro) por cm². Por síntesis *in situ* mediante desprotección fotoquímica, se alcanzan fácilmente los 65.000 puntos (de 50 micrometros de diámetro) por cm².

15

Los microchips de ADN pueden aplicarse a estudios de expresión génica, resecuenciación de genomas y genotipado, principalmente. La expresión de ARN puede analizarse para miles de genes usando muestras de tejidos enfermos (cáncer, infecciones virales, infecciones bacterianas, etc.) o usando los agentes infecciosos propiamente dichos (virus, bacterias, hongos, etc.). El descubrimiento de los genes implicados en estos procesos permite desarrollar y diseñar nuevas drogas, nuevas técnicas de diagnóstico, etc. Estudios de resecuenciación y genotipado pueden usarse para descubrir mutaciones y polimorfismos de un nucleótido (SNP).

30

Se han descrito diversas estrategias para la detección de SNPs. Una de ellas es la mencionada más arriba (Hacia *et al.*, 1998); la estrategia de secuenciación por hibridación a un microchip de octa y decanucleótidos amplificados mediante la unión de pentanucleótidos adyacentes (Parinov *et al.*, 1996); las nuevas estrategias que adaptan tratamientos enzimáticos como la desarrollada por Gerry *et al.* (1999), que combina la reacción en cadena de la polimerasa (PCR) y la reacción de detección de la ligasa (LDR) con un chip de código universal y que permite detectar mutaciones en genomas humanos que están presentes en proporciones inferiores al 1% respecto de la copia del ADN salvaje. Otro método utiliza la actividad ADN polimerasa del fragmento Klenow de la ADN polimerasa de *Escherichia coli* para elongar un oligonucleótido hibridado a otro que a su vez sirve de molde para extender el primero (Hacia, 1999). Los oligonucleótidos se pueden diseñar de tal manera que por cada cuatro pares de oligonucleótidos la posición interrogante corresponda a la primera del molde que ha de ser copiada (A, C, G o T), y si la reacción de extensión se hace en presencia de los cuatro dideoxinucleótidos (ddNTPs) marcados con un compuesto fluorescente distinto, entonces es posible averiguar por el tipo de fluorescencia cuál es el nucleótido en la posición interrogante. Siguiendo esta metodología, es posible sintetizar *in situ* oligonucleótidos de doble cadena que a su vez puedan ser empleados para el estudio de interacciones proteína-ADN (Bulyk *et al.*, 1999), y por tanto abre un nuevo abanico de posibilidades para el descubrimiento de nuevas técnicas de diagnóstico y nuevos fármacos.

45

Otro campo de aplicación de los microchips es el de la identificación de especies de microorganismos, principalmente de las variantes o cepas (más o menos virulentas) de la misma especie (Gingeras *et al.*, 1998), bien para aplicaciones tradicionales (resistencia a drogas, toxinas, factores de patogenicidad, etc.) o bien para aplicaciones ecológicas (biodiversidad, dispersión polimórfica, etc.). Gingeras *et al.* construyeron un chip de ADN con oligonucleótidos interrogando todas las posiciones (en las dos cadenas) de un fragmento de ADN de 705 pb del gen *rpoB* de *Mycobacterium tuberculosis*, para analizar en una colección de 63 aislados clínicos de *M. tuberculosis* la existencia de mutaciones que confieren resistencia a rifampicina. La identificación de especies se basó en la existencia de polimorfismos específicos de especie que pueden ser fácilmente determinados con un microchip de ADN.

50

Otro ejemplo de la utilización de microchips de ADN para identificar bacterias se describe en la patente estadounidense número 5.925.522, donde Wong *et al.* describen métodos para la detección de *Salmonella* mediante chips de ADN con secuencias de oligonucleótidos específicas.

55

La tecnología de chips de ADN se ha empezado a desarrollar recientemente (en 1996) y avanza vertiginosamente. Esto permite detectar cada vez mejor y con mayor fiabilidad alteraciones genéticas en mezclas complejas que hasta ahora sólo era posible mediante métodos muy laboriosos y de larga duración. Un microchip de ADN adecuadamente diseñado permite identificar mutaciones puntuales en pocas horas, mientras que se necesitarían días, y a veces semanas, para hacerlo por la metodología convencional.

60

La tecnología de chips de ADN se ha empezado a desarrollar recientemente (en 1996) y avanza vertiginosamente. Esto permite detectar cada vez mejor y con mayor fiabilidad alteraciones genéticas en mezclas complejas que hasta ahora sólo era posible mediante métodos muy laboriosos y de larga duración. Un microchip de ADN adecuadamente diseñado permite identificar mutaciones puntuales en pocas horas, mientras que se necesitarían días, y a veces semanas, para hacerlo por la metodología convencional.

65

Descripción**Descripción breve**

5 La invención se relaciona con un método para diseñar una terapia antiviral individual para un sujeto basado en la detección de genomas memoria minoritarios en una cuasi-especie viral responsable de la resistencia del virus a un fármaco o combinación de fármacos.

10 La solución proporcionada por esta invención se basa en el descubrimiento de la existencia de genomas memoria virales que, aun siendo genomas minoritarios, reflejan la historia evolutiva del virus en el organismo infectado, y comprende el análisis de cuasi-especies virales, mediante cualquier técnica apropiada, por ejemplo, microchips de ADN, ensayo de la huella del heteroduplex (HTA) y clonaje molecular.

15 El método proporcionado por esta invención permite, entre otras aplicaciones, detectar e identificar genomas memoria minoritarios; analizar cuasi-especies virales responsables de la resistencia del virus a fármacos, o relevantes a la selección frente a los sistemas de defensa (respuesta inmune) del organismo infectado y al diseño de regímenes de terapias personalizadas.

20 El Ejemplo 1 describe la detección y caracterización de genomas minoritarios del virus de la fiebre aftosa (VFA) mediante microchips de ADN; el Ejemplo 2 ilustra la detección y caracterización de genomas minoritarios del virus de la inmunodeficiencia humana (VIH) portadores de mutaciones de resistencia a zidovudina (también llamada azidotimidina, AZT) mediante microchips de ADN; el Ejemplo 3 describe la detección y caracterización de genomas memoria en poblaciones de VFA; el Ejemplo 4 ilustra la detección de genomas memoria en sujetos de VIH portadoras de mutaciones que confieren resistencia a fármacos en enfermos tratados; el Ejemplo 5 describe la detección de genomas memoria en cuasi-especies del virus de la hepatitis C (VHC); y el Ejemplo 6 ilustra la detección de genomas memoria del virus de la hepatitis B (VHB) portadores de mutaciones que confieren resistencia a fármacos en enfermos tratados.

Descripción detallada de la invención**Definiciones**

30 El término “cuasi-especie” se refiere a una población viral constituida por distribuciones dinámicas de genomas mutantes y recombinantes no idénticos pero estrechamente relacionados entre sí, que están sujetos a un proceso continuo de variación genética, competición y selección, y que actúan como una unidad de selección.

40 El término “genoma memoria” de una cuasi-especie se refiere a la capacidad de la cuasi-especie para mantener por encima del umbral de fondo genomas mutantes específicos que, alguna vez a lo largo de la historia de la cuasi-especie en un organismo infectado, constituyó la secuencia mayoritaria o el promedio de la cuasi-especie.

Se define como “umbral de fondo” la proporción con que aparecen en la cuasi-especie los genomas distintos del genoma mayoritario y de genomas memoria. Tal proporción es de aproximadamente 10^{-5} (0,001%) en la cuasi-especie.

45 Se define como “genoma mayoritario o promedio” el genoma viral con la secuencia de nucleótidos que representa una proporción igual o superior al 50% de una cuasi-especie.

50 Se define “genoma tipo” o “virus tipo” como el genoma de un aislado viral con una secuencia que está considerada por la comunidad científica como de tipo salvaje, y que sirve como genoma de referencia para esa especie. En el caso del VIH-1, como genoma salvaje se considera la cepa HXB2 (número de acceso en la base de datos GenBank: K03455).

55 Se define como “genoma minoritario” el genoma viral con una secuencia de nucleótidos que está presente en menos del 50% de la cuasi-especie. Un genoma que en una cuasi-especie es mayoritario, puede ser minoritario en otra cuasi-especie; y viceversa.

Se denomina “genoma memoria” a la secuencia de nucleótidos de un genoma viral que presenta al menos una mutación respecto del genoma mayoritario de una cuasi-especie determinada y que proporciona información sobre la historia evolutiva de dicha cuasi-especie.

60 Tal como se utiliza en esta descripción, “genoma memoria minoritario” se refiere a la secuencia de nucleótidos de un genoma viral que presenta al menos una mutación respecto del genoma mayoritario en una cuasi-especie determinada, proporciona información sobre la historia evolutiva de dicha cuasi-especie, y se encuentra en menos del 50%.

65 El término “ácido nucleico” se refiere a una secuencia de ácido desoxirribonucleico, ácido ribonucleico o ácido nucleico peptídico, de longitud igual o superior a 4 nucleótidos (abreviado como nt) que puede ser de cadena sencilla o cadena doble.

ES 2 273 835 T3

Se define como “oligonucleótido” una molécula de ADN de banda sencilla de entre 4 y 250 nt de longitud, que puede hibridar con una secuencia de ADN complementaria.

5 El término “posición nucleotídica” se refiere al lugar que ocupa cada uno de los nucleótidos en una secuencia de ácidos nucleicos.

“Secuencia codificante” es la secuencia de nucleótidos que especifica la secuencia de aminoácidos de una proteína.

10 Un “codón” o “tripleto” es una secuencia de tres nucleótidos consecutivos que dentro de una secuencia codificante especifica un aminoácido. Por ejemplo, el tripleto ATG especifica o codifica el aminoácido metionina (M).

15 Una “mutación” es una alteración de la secuencia de nucleótidos de un ácido nucleico respecto a otra secuencia de referencia. Dicha alteración puede ser una sustitución de un nucleótido por otro, una inserción o una delección de uno o más nucleótidos.

20 En la presente descripción se define “oligonucleótido memoria (OM) de una cuasiespecie” como un ácido nucleico de una longitud desde 4 a 250 nucleótidos y que es igual o complementario a una secuencia del genoma viral mayoritario o promedio, excepto en al menos un nucleótido. Un nucleótido memoria puede ser también un nucleótido mayoritario en otra cuasiespecie.

25 “Oligonucleótido memoria de tipo 1 (OM1)” se define como un ácido nucleico de una longitud desde 4 a 250 nucleótidos que es igual y/o complementario a un genoma viral mayoritario o promedio excepto en 1 a 6 posiciones centrales.

30 “Oligonucleótido memoria de tipo 2 (OM2)” se refiere a ácidos nucleicos con una longitud de 5-50 nt que se forman por apilamiento de dos oligonucleótidos tras hibridar con otro ácido nucleico complementario del virus. Uno de los oligonucleótidos apilados está formado por una combinación de cuatro oligonucleótidos que difieren (es decir, tienen una posición interrogante) en la posición inmediatamente adyacente al oligonucleótido anterior, y que llevan un marcador fluorescente diferente unido covalentemente al otro extremo.

35 “Apilamiento de bases” o “apilamiento de nucleótidos” se refiere a la interacción establecida entre dos bases adyacentes o los extremos de dos oligonucleótidos mediante apilamiento de sus purinas o pirimidinas. En este tipo de interacción, no se forman enlaces covalentes entre las bases adyacentes, pero tal interacción incrementa la estabilidad de la unión de ambos nucleótidos a la cadena complementaria común a ambos. De esta forma, por ejemplo, un oligonucleótido de 5 nt puede permanecer establemente unido a su cadena complementaria únicamente si se produce apilamiento con un oligonucleótido adyacente de 5 nt o más.

40 “Posición interrogada” es la posición nucleotídica de la secuencia del genoma viral para la cual la composición es desconocida.

“Posición interrogante o discriminatoria” es la posición nucleotídica del oligonucleótido memoria ocupada por uno de los cuatro posibles nucleótidos (A, C, G, T).

45 “Oligonucleótido memoria de tipo 3 (OM3)” es un ácido nucleico de 5 a 250 nt que consta de dos partes, la sección 5’ del oligonucleótido es complementaria a otro oligonucleótido que está ausente del genoma viral, y la sección 3’ es complementaria al genoma viral, siendo la última posición una posición interrogante.

50 “Oligonucleótido memoria de tipo 4 (OM4)” es un ácido nucleico de 5 a 250 nt de longitud complementario al genoma viral que lleva una sustancia fluorescente unida covalentemente al extremo 3’. Un nucleótido OM3 se puede usar junto con una OM4 para detectar mutaciones por la técnica de PCR/LDR descrita por Gerry *et al.* (1999), utilizando un microchip de ADN con oligonucleótidos complementarios a la parte 5’ de OM3.

55 “Oligonucleótido memoria de tipo 5 (OM5)” es un ácido nucleico de 5-250 nt de longitud complementario al genoma viral en la cual la última posición en el extremo 3’ es anterior a una posición interrogada del genoma viral.

60 “Oligonucleótido memoria de tipo 6 (OM6)” es un ácido nucleico de 5 a 250 nucleótidos complementarios a una secuencia de un genoma mayoritario de una cuasiespecie viral con inserciones de 1 a 10 nucleótidos respecto a dicha secuencia del genoma mayoritario.

65 “Oligonucleótido memoria de tipo 7 (OM7)” es un ácido nucleico de 5 a 250 nucleótidos complementarios a una secuencia de un genoma mayoritario de la cuasiespecie viral con delecciones de 1-10 nucleótidos respecto a dicha secuencia del genoma mayoritario.

“Oligonucleótido memoria de tipo 8 (OM8)” es un ácido nucleico de 5 a 250 nucleótidos complementario a una secuencia mutante previamente descrita en las bases de datos.

ES 2 273 835 T3

El término “sonda” se refiere a ácidos nucleicos de 5 a 250 pares de bases de longitud comprendidos por secuencias de nucleótidos específicas que permiten una hibridación total o parcial con secuencias diana complementarias bajo condiciones definidas.

5 Secuencias “diana” son secuencias de ácidos nucleicos susceptibles de hibridar con las sondas de oligonucleótidos. En la presente invención las secuencias diana se marcan con un compuesto radiactivo o fluorescente cuando se usan las técnicas de microchips de ADN, pero no cuando se hace por ensayo de la huella del heteroduplex (HTA) (Gerotto *et al.*, 1999).

10 “Secuencias flanqueantes” en un oligonucleótido son aquellas (de 5 a 100 nt) que acompañan a la/s posición/es interrogada/s y que, juntas, permiten la hibridación del oligonucleótido a los genomas salvaje o memoria (mutante).

15 “Oligonucleótido salvaje” es aquél que posee total identidad de secuencia con el genoma salvaje o genoma correspondiente.

15 “Oligonucleótido mutante” es aquél que posee total identidad de secuencia con el genoma salvaje correspondiente, salvo en la/s posición/es interrogante/s, en las cuales posee la secuencia correspondiente a una cepa mutante en esa posición.

20 El término “hibridación” se refiere a un proceso mediante el cual, en ciertas condiciones, dos cadenas complementarias de ácidos nucleicos se unen de forma antiparalela mediante la formación de puentes de hidrógeno para formar ácidos nucleicos de cadena doble, según las reglas de apareamiento entre bases nitrogenadas.

25 “Hibridación total” o “hibridación 100%” se refiere a la hibridación que tiene lugar cuando todos los nucleótidos de una sonda u oligonucleótido aparecen con una secuencia diana, o viceversa.

Un “desapareamiento” ocurre cuando en al menos una posición de un ácido nucleico de doble cadena, ambas cadenas poseen dos nucleótidos no complementarios.

30 Un “híbrido” es el resultado del proceso de hibridación entre dos ácidos nucleicos de cadena sencilla. Los desapareamientos en las posiciones centrales tienen un efecto desestabilizador mayor que cuando se localizan en los extremos. La estabilidad depende de la longitud, el número de desapareamientos, y los factores externos como la temperatura, la fuerza iónica del medio y el pH. Cuanto mayor sea la longitud, menor el número de desapareamientos, menor la fuerza iónica, menor temperatura y pH cercano a la neutralidad, mayor será la estabilidad del híbrido.

35 “PCR anidada” es un método de amplificación enzimática del ADN que consiste en la realización de dos rondas sucesivas de PCR, la segunda con una pareja de oligonucleótidos interiores a los empleados en la primera PCR. Mediante estas dos rondas de PCR se logra amplificar cantidades extremadamente pequeñas del ADN de partida, lo que resulta muy útil en el caso de muestras clínicas, en las que el virus se encuentra en muy bajas cantidades.

40 “Una muestra sospechosa de contener la cuasiespecie viral” es cualquier muestra procedente de un animal, una planta, una bacteria o un cultivo celular que puede estar infectada con al menos una cuasiespecie viral.

45 *Método para la detección de genomas minoritarios*

La invención emplea una técnica para la detección de genomas minoritarios de una población de ácidos nucleicos de una cuasiespecie viral, presentes en menos del 50%, y conteniendo, al menos, una mutación respecto al genoma mayoritario de dicha cuasiespecie, en adelante, método de la invención, que comprende:

50 a) extraer el ácido nucleico de dicha cuasiespecie viral a partir de una muestra sospechosa de contener dicha cuasiespecie viral;

 b) amplificar al menos un fragmento del ácido nucleico de dicha cuasiespecie viral; y

55 c) detectar y analizar la existencia de genomas minoritarios mediante el empleo de técnicas basadas en microchips de ADN, el ensayo de la huella del heteroduplex (HTA) y el clonaje molecular.

Alternativa 1

60 En una realización particular, la invención proporciona un método para la detección de genomas minoritarios que incluye el empleo de microchips de ADN. De forma más concreta, la invención proporciona una técnica para la detección de genomas minoritarios de una población de ácidos nucleicos de una cuasiespecie viral, presentes en una proporción inferior al 50%, y que contiene, al menos, una mutación respecto al genoma mayoritario de dicha cuasiespecie. El método consiste en las siguientes etapas:

65 a) extraer el ácido nucleico de dicha cuasiespecie viral a partir de una muestra sospechosa de contener dicha cuasiespecie viral;

ES 2 273 835 T3

- b) amplificar al menos un fragmento del ácido nucleico de dicha cuasiespecie viral;
- c) marcar el fragmento o fragmentos amplificados con un compuesto marcador;
- 5 d) construir un microchip de ADN de manera que se generen puntos que comprenden:
- i) al menos, un oligonucleótido que puede servir como un control positivo;
 - 10 ii) al menos, un oligonucleótido que puede servir como un control negativo;
 - iii) al menos, un oligonucleótido memoria; y
 - iv) unos medios que puedan usarse para realizar una curva de calibrado;
- 15 e) poner en contacto dichos fragmentos amplificados en la etapa b) y marcados en la etapa c) con los oligonucleótidos presentes en el microchip de ADN preparado en la etapa d) bajo condiciones que permiten la hibridación, únicamente cuando todos los nucleótidos de un oligonucleótido presente en dicho microchip de ADN aparean con una secuencia de nucleótidos presente en dichos fragmentos amplificados y marcados;
- 20 f) identificar los oligonucleótidos presentes en dicho microchip de ADN que han hibridado con dichos fragmentos amplificados y marcados, descartando las hibridaciones negativas o ruido de fondo; y
- g) seleccionar los oligonucleótidos presentes en dicho microchip de ADN que han hibridado con dichos fragmentos amplificados y marcados, y que por interpolación con la curva de calibrado muestran una proporción de dichos fragmentos en el ADN de la cuasiespecie más bajos del 50% característico de genomas minoritarios.

En una realización particular, dichos genomas minoritarios son genomas memoria minoritarios, que pueden estar presentes en la cuasiespecie viral en una proporción comprendida entre el 0,1 y el 10% de la cuasiespecie.

30 La cuasiespecie viral puede corresponder a un virus con genoma de ADN o un virus con genoma de ARN. En una realización particular, dicha cuasiespecie viral pertenece a un virus seleccionado del grupo formado por el virus de la inmunodeficiencia humana tipo 1 (VIH-1), el virus de la inmunodeficiencia humana tipo 2 (VIH-2), el virus de la hepatitis C (VHC), y el virus de la hepatitis B (VHB).

35 La técnica de la invención comienza con la extracción del ácido nucleico de la cuasiespecie viral a partir de una muestra sospechosa de contener dicha cuasiespecie viral, por ejemplo, una muestra una muestra clínica o un muestra seleccionada de un cultivo viral. La extracción del ácido nucleico se realiza mediante técnicas convencionales (Sambrook *et al.*, 1989).

40 La amplificación del fragmento o fragmentos del ácido nucleico extraído de la cuasiespecie viral se puede realizar por cualquier método convencional. En una realización particular, dicha amplificación se realiza mediante técnicas enzimáticas, por ejemplo, mediante la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR) o la amplificación basada en transcripción (TAS). En caso de que el ácido nucleico extraído sea un ARN, se realiza una transcripción inversa (RT), mediante técnicas convencionales, del ARN viral previa a la etapa b) de amplificación. En una realización práctica de la técnica descrita aquí, la amplificación del fragmento o fragmentos de ácido nucleico extraído de la cuasiespecie viral se realiza por RT-PCR (en caso de partir de virus ARN no integrados en el ADN de la célula hospedadora), o por PCR sencilla o PCR anidada (para virus ADN o en el caso de virus ARN integrados en el genoma de las células hospedadoras). Se amplifica el mismo fragmento a partir de un aislado del virus tipo (cepa salvaje).

50 Aunque el fragmento o fragmentos del ácido nucleico extraído de la cuasiespecie viral a amplificar puede proceder de cualquier gen o región del genoma viral, preferentemente, dicho fragmento o fragmentos proceden de la totalidad, o una parte, de, al menos, un gen imprescindible para la replicación o la persistencia del virus en el organismo infectado. En una realización particular, dicho fragmento o fragmentos proceden de un gen imprescindible seleccionado del grupo formado por el fragmento de la proteasa (PR) del gen *pol* de VIH, el fragmento de la transcriptasa inversa (RT) del pool de genes de VIH, el fragmento de la integrasa del pool de genes de VIH, el gen *env* de VIH, el gen *gag* de VIH, el gen de la proteína no estructural NS5A de VHC, la región comprendida entre los nucleótidos 175-215 del VHC, la región comprendida entre los nucleótidos 310-350 del VHC, y el fragmento de la transcriptasa inversa (RT) del gen *pol* de VHB.

60 Los fragmentos amplificados, en general, se purifican y se marcan con un compuesto marcador apropiado, tal como, una sustancia radiactiva, por ejemplo, ³²P, ³³P, etc., un compuesto fluorescente, por ejemplo Cy3, Cy5, etc., o un compuesto detectable por reacción colorimétrica, por ejemplo un compuesto que da lugar a una reacción enzimática coloreada. En general, se marcan dos fragmentos de ADN diferentes; uno cuya secuencia es exactamente idéntica a la del virus salvaje, y otro comprendido por la combinación de fragmentos de secuencia diferentes procedente de la cuasiespecie a estudiar.

ES 2 273 835 T3

En una realización particular, los fragmentos de ácido nucleico amplificados, por ejemplo, por RT-PCR o por PCR, a partir del ARN o ADN viral, se marcan con precursores nucleotídicos que llevan un fluorocromo, por ejemplo, Cy3 o Cy5-dCTP, o Cy3 o Cy-dUTP, bien incluyéndolo en los oligonucleótidos utilizados como cebadores en la reacción de amplificación por RT-PCR, bien por marcaje mediante extensión de hexanucleótidos al azar, o bien por marcaje químico con reactivos intercalantes como el psoraleno-biotina. En una realización preferida, como se mencionó previamente, es conveniente marcar dos muestras, una conteniendo la secuencia salvaje marcada con un fluorocromo como control (muestra de referencia), y otra conteniendo las secuencias a estudiar marcadas con otro fluorocromo (muestra test). La posibilidad de marcar con fluorocromos distintos permite hacer hibridaciones con las dos muestras a la vez en un mismo microchip.

10

A continuación, se construye un microchip de ADN con puntos que comprenden:

- i) al menos, un oligonucleótido que sirve como un control positivo;
- 15 ii) al menos, un oligonucleótido que sirve como un control negativo;
- iii) al menos, un oligonucleótido memoria; y
- 20 iv) unos medios que pueden permitir realizar una curva de calibrado.

25

Los oligonucleótidos usados como controles permiten evaluar la calidad de la hibridación o la cuantificación de la misma, con el fin de poder correlacionar la intensidad de la señal de hibridación con la abundancia de una mutación en la cuasiespecie viral.

25 Como control positivo puede utilizarse un oligonucleótido cuya secuencia es 100% complementaria a una zona de secuencia conocida del genoma mayoritario o del genoma del virus salvaje.

El control negativo puede seleccionarse del grupo formado por: i) un oligonucleótido cuya secuencia es complementaria a una zona de secuencia conocida del genoma mayoritario del virus salvaje, excepto en al menos una posición; 30 ii) un oligonucleótido cuya secuencia es complementaria a una zona de secuencia conocida del genoma mayoritario o del genoma salvaje, excepto en la posición interrogante; y iii) un oligonucleótido cuya secuencia es complementaria a una secuencia conocida del genoma mayoritario o del genoma del virus salvaje excepto en la posición interrogante y en, al menos, un nucleósido que flanquea la posición interrogante. El oligonucleótido memoria se selecciona del grupo comprendido por los oligonucleótidos memoria identificados como OM1, OM2, OM3, OM4, OM5, OM6, OM7 y 35 OM8 y combinaciones de estos.

La curva de calibrado está formada por series de combinaciones de nucleótidos en proporciones variables y conocidas, siendo uno de ellos 100% complementario a una zona de secuencia conocida del genoma mayoritario o del genoma de salvaje, y siendo el otro un oligonucleótido que difiere del anterior en, al menos, una posición. En una 40 realización particular, dicha curva de calibrado se elabora a partir de una serie de mezclas de oligonucleótidos en proporciones variables y conocidas, siendo uno de ellos 100% complementario a una zona de secuencia conocida del genoma mayoritario o del genoma salvaje, y siendo el otro un oligonucleótido que difiere del anterior en la posición interrogante.

45 El microchip de ADN se puede construir mediante técnicas convencionales. En una realización particular, dicho microchip de ADN está compuesto por puntos de oligonucleótidos previamente sintetizados, mientras que en otra realización particular, dicho microchip de ADN está compuesto por puntos de oligonucleótidos sintetizados *in situ*. En una realización particular, se construyen microchips de ADN que contienen oligonucleótidos memoria que interrogan cada una de las posiciones del genoma viral que se desean estudiar. Es decir, cada posición interrogada del genoma viral está representada por cuatro puntos diferentes en el microchip, cada uno de ellos conteniendo un oligonucleótido interrogante por cada base A, C, G o T. También se incluyen otros tipos de oligonucleótidos que sirven como controles de la 50 calidad de la hibridación o cuantificación de la misma, para poder correlacionar intensidad de la señal de hibridación con la abundancia de una mutación en la cuasiespecie viral. En otra realización particular, se contempla la posibilidad de construir microchips de ADN con oligonucleótidos memoria con más de una posición interrogante (2 ó 3) por codón analizado. De la misma forma, se pueden incluir oligonucleótidos memoria interrogando por inserciones y deleciones. 55

Los fragmentos de ADN amplificados en la etapa b) y marcados en la etapa c) (ácidos nucleicos diana) se ponen en contacto con los oligonucleótidos presentes en el microchip de ADN preparado en la etapa d) bajo condiciones que permiten la hibridación, únicamente, cuando todos los nucleótidos de un oligonucleótido presente en dicho microchip 60 de ADN aparean con una secuencia de nucleótidos presente en dichos fragmentos amplificados y marcados, es decir, en condiciones en las que únicamente hibridan secuencias de oligonucleótidos complementarias con el 100% de las secuencias de los fragmentos marcados y amplificados. La selección de las condiciones de hibridación adecuadas depende de diversos factores, que incluye el tamaño de los oligonucleótidos, y que éstos puedan ser fijados en cada caso, fácilmente, por un técnico de laboratorio especializado.

65

Una vez que la hibridación está completa, debe confirmarse que ésta ha tenido lugar y que se identificaron los oligonucleótidos presentes en el microchip de ADN, descartando las hibridaciones negativas o el ruido de fondo. En una realización particular, si el microchip se ha construido con OM1, OM2 o con oligonucleótidos complementarios a la

ES 2 273 835 T3

parte 5' de OM3, tras la hibridación se lava con el tampón adecuado, en unas condiciones apropiadas, y se comprueba el resultado de la hibridación mediante el escaneado del microchip de ADN con un escáner provisto de un microscopio confocal y al menos dos láseres que emiten luz a diferente longitud de onda y los filtros correspondientes a los fluorocromos empleados para marcar los fragmentos de ADN amplificados, y un soporte informático que permite obtener la imagen computerizada del resultado de la hibridación. Ciertos programas permiten cuantificar la intensidad de la hibridación y hacer curvas de calibrado usando los controles de concentraciones conocidas incluidos en el microchip de ADN.

Finalmente, se interpretan los resultados. El patrón de hibridación obtenido indica la presencia o ausencia de genomas minoritarios, por ejemplo, genomas memoria minoritarios en la cuasiespecie viral para cada una de las posiciones interrogadas. Por lo tanto, en el análisis muestras clínicas es posible determinar:

i) si en la cuasiespecie viral existen genomas con mutaciones nucleotídicas en los codones implicados en resistencia a fármacos antivirales. La presencia de dichas mutaciones en el genoma mayoritario dará una indicación del patrón actual de resistencias, generalmente derivado de la terapia antiviral a que el paciente está recibiendo en el momento del análisis. La presencia de dichas mutaciones en genomas memoria minoritarios resultaría a partir de la historia previa de tratamiento antiviral del paciente. En cualquier caso, la presencia de mutaciones de resistencia en genomas que están más o menos bien representados (mayoritario o memoria) dentro de la cuasiespecie viral producirá una falta de respuesta al fármaco correspondiente a partir de ese momento. Por ello, tal fármaco deberá ser excluido (como monoterapia o en combinación de dos o más fármacos) en protocolos terapéuticos futuros diseñados por el médico o veterinario con objeto de suprimir o reducir la replicación viral; y

ii) La presencia en la cuasiespecie viral de genomas (mayoritarios o minoritarios) con mutaciones nucleotídicas en los codones implicados en el escape del sistema inmune; de esta forma, el médico o veterinario puede tomar las medidas oportunas relacionadas con el uso de anticuerpos específicos, vacunas u otros tratamientos, la acción de los cuales está basada en sus efectos sobre el sistema inmunitario.

La invención contempla la posibilidad de cuantificar los genomas minoritarios, en particular, los genomas memoria minoritarios. Para ello,

(1) se diseñó un microchip que tuviera:

a) controles de oligonucleótidos memoria y oligonucleótidos salvaje conocidos por duplicado;

b) combinaciones de oligonucleótidos memoria y oligonucleótidos salvaje: a) en distintas proporciones, mutante/salvaje, por ejemplo, 10^{-5} , 10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 1, 10, 10^2 , 10^3 , 10^4 , 10^5 y

(2) La hibridación del microchip se hace con:

a) la muestra test y la muestra test de referencia (fragmento de la misma longitud y de secuencia igual a la salvaje) marcadas con distinto fluorocromo. La cuantificación se hace mediante diferencias en la intensidad de la hibridación con el juego de oligonucleótidos mencionados previamente en (1)b); y

b) la muestra test y una mezcla formada por fragmentos puros amplificados de la secuencia salvaje y una secuencia memoria conocida en distintas proporciones y marcadas con el mismo fluorocromo que es diferente al usado en la muestra test. Para realizar una curva de cuantificación es necesario llevar a cabo el proceso de hibridación tantas veces como concentraciones haya que estudiar.

Los microchip lavados pueden leerse usando un escáner provisto de un microscopio confocal y dos láseres que emiten luz con diferente longitud de onda. De este modo, es posible leer simultáneamente dos fluorocromos en un mismo microchip.

La determinación de la intensidad de hibridación en los oligonucleótidos minoritarios o memoria por comparación con controles, identifica inmediatamente las mutaciones que pueden formar parte de un genoma minoritario o memoria.

Alternativa 2

En otra realización particular, la invención proporciona un método para la detección de genomas minoritarios que incluye el empleo de técnicas basadas en el ensayo de la huella del heterodúplex [HTA] (Gerotto, *et al.*, 1999). Esta alternativa del método de la invención puede ser empleada para el seguimiento de cuasiespecies virales en un mismo paciente. De forma más concreta, la invención proporciona un método para la detección de genomas virales minoritarios en una población de ácidos nucleicos de una cuasiespecie viral, presentes en una proporción de menos del 50% y conteniendo, al menos, una mutación respecto al genoma mayoritario de dicha cuasiespecie, que consiste en:

a) extraer el ácido nucleico de dicha cuasiespecie viral a partir de una muestra sospechosa de contener dicha cuasiespecie viral;

ES 2 273 835 T3

b) amplificar al menos un fragmento del ácido nucleico de dicha cuasiespecie viral;

c) clonar los fragmentos de ADN amplificados en b) en un vector adecuado;

5 d) determinar la secuencia del genoma mayoritario de una cuasiespecie viral para el fragmento amplificado;

e) amplificar los fragmentos de ADN clonados en la etapa c) y marcar los fragmentos amplificados con un compuesto marcador;

10 f) poner en contacto, en una reacción de hibridación, los fragmentos amplificados y marcados procedentes de la etapa e) con los fragmentos amplificados directamente del ácido nucleico de la cuasiespecie viral procedentes de la etapa b); y

15 g) resolver las diferentes secuencias virales e identificar las mutaciones indicativas de los genomas minoritarios presentes en la cuasiespecie viral.

La extracción del ácido nucleico y la amplificación del fragmento o fragmentos del ácido nucleico de dicha cuasiespecie viral se realiza como se mencionó previamente en la método alternativo de la invención que emplea de microchips de ADN.

20 El clonaje de los fragmentos de ácido nucleico y la secuenciación de los fragmentos y secuencias se puede realizar mediante técnicas convencionales conocidos por los técnicos de laboratorio especializados. En una realización particular, el clonaje de los fragmentos de ADN amplificados en la etapa b) se realizó en un plásmido de alto número de copias. Información sobre técnicas de clonaje y secuenciación de secuencias de ácidos nucleicos puede encontrarse en Sambrook *et al.*, 1989.

El compuesto marcador utilizado en la etapa e) para marcar los fragmentos de ADN amplificados y clonados en la etapa c) puede ser cualquier marcador adecuado. En una realización particular, se marcan los extremos 5' de dichos fragmentos de ADN con polinucleótido quinasa y [γ -³²P]-ATP.

30 Tras la reacción de hibridación, la resolución de las diferentes secuencias virales se puede realizar mediante:

g.i) fraccionamiento de los híbridos formados en la etapa f) mediante electroforesis en geles de poliacrilamida en condiciones no desnaturalizantes;

35 g.ii) identificar la existencia de genomas minoritarios por el número de mutaciones en relación al salto repentino en la movilidad electroforética;

40 g.iii) extraer los ADN hibridados y fraccionados mediante electroforesis en gel de poliacrilamida por elución;

g.iv) amplificar los fragmentos eluidos en la etapa g.iii);

g.v) secuenciar los fragmentos amplificados en la etapa g.iv); y

45 g.vi) comparar las secuencias deducidas en la etapa g.v) e identificar las mutaciones indicativas de los genomas minoritarios presentes en la cuasiespecie viral.

Según esta alternativa, la existencia de cambios de nucleótidos entre la sonda marcada y el ADN diana se manifiesta por un retraso en la movilidad electroforética de los heterodúplex con desapareamientos formados. Como control de migración del homodúplex, la sonda se hibrida con sus propias secuencias sin marcar, de manera que la diferencia de migración con los heterodúplex es proporcional al número de cambios de nucleótidos. De esta forma es posible detectar la presencia de genomas memoria y calcular el número de cambios nucleotídicos en la zona estudiada para cada uno. Un microchip de ADN que interroga todas las posiciones de los fragmentos de ADN estudiados dará la identidad de cada nucleótido.

55 Alternativa 3

En otra realización particular, la invención proporciona un método para la detección de genomas minoritarios que incluye la determinación de la secuencia consenso de la cuasiespecie viral o el clonaje molecular y posterior secuenciación de otros clones obtenidos. De forma más concreta, la invención proporciona un método para la detección de genomas minoritarios de una población de ácidos nucleicos de una cuasiespecie viral, presentes en una proporción inferior al 50%, y conteniendo, al menos, una mutación respecto al genoma mayoritario de dicha cuasiespecie, que comprende:

65 a) extraer el ácido nucleico de dicha cuasiespecie viral a partir de una muestra sospechosa de contener dicha cuasiespecie viral;

b) amplificar al menos un fragmento del ácido nucleico de dicha cuasiespecie viral;

ES 2 273 835 T3

c) determinar la secuencia del genoma mayoritario de la cuasiespecie viral para dicho fragmento amplificado;

d) opcionalmente, clonar el fragmento de ácido nucleico amplificado en un vector;

5 e) secuenciar el fragmento clonado; y

f) comparar las secuencias deducidas en las etapas c) y e) e identificar las mutaciones indicativas de los genomas minoritarios presentes en la cuasiespecie viral.

10 La extracción del ácido nucleico y la amplificación del fragmento o fragmentos del ácido nucleico de dicha cuasiespecie viral se realiza de forma similar al procedimiento descrito previamente en relación a la alternativa del método de la invención que concierne al empleo de microchips de ADN.

15 El clonaje de los fragmentos de ácido nucleico y la secuenciación de los fragmentos y secuencias se puede realizar mediante métodos convencionales conocidos por los técnicos expertos en la materia. Información sobre técnicas de clonaje y secuenciación de secuencias puede encontrarse en Sambrook *et al.*, 1989.

20 La secuencia del genoma viral mayoritario y su posterior comparación por alineamiento de las secuencias, por ejemplo, mediante el empleo de PILEUP o CLUSTAL, dará la información precisa sobre la existencia de genomas minoritarios y las mutaciones exactas que les caracterizan.

25 Un aspecto preferente de la invención concierne a la utilización de la información de la existencia de genomas memoria minoritarios virales para diseñar nuevas terapias antivirales personalizadas. Como se ha indicado previamente, la presencia de un genoma memoria portador de una mutación que, por datos bibliográficos o estudios anteriores, se sabe que está asociada con la resistencia a un medicamento, implicaría la recomendación de no utilizar este fármaco en la terapia que, a partir de ese momento, requiera el paciente. La continuación en el uso de un fármaco o fármacos para los que una cuasiespecie viral ha desarrollado mutaciones de resistencia impediría la actuación del fármaco antiviral, que no obstante, continuaría produciendo efectos secundarios, de variables grados de severidad, en el paciente. Por lo tanto, la adecuada determinación del perfil de mutaciones de resistencia a fármacos antivirales produce una mejora de la calidad de vida del paciente (por la eliminación de los efectos secundarios de uno de los fármacos a los que el virus ha desarrollado resistencia), y la importante reducción económica para el Sistema de Salud que normalmente corre con el coste de la medicación (del fármaco o fármacos que, a pesar de ser inactivos frente al virus, podrían seguir siendo administrados en el caso de desconocer el perfil de resistencias del virus a este/estos fármacos). En este contexto, una de las principales aplicaciones prácticas de la invención radica en el incremento considerable en la cantidad de información disponible sobre mutaciones de resistencia para cada paciente individual ya que, además de las mutaciones presentes en el genoma mayoritario (las únicas que se determinan mediante las técnicas convencionales actuales), se suma la información acerca de los genomas memoria minoritarios en la cuasiespecie viral, que igualmente reduce (a corto o medio plazo) la eficacia del fármaco.

40 Otra aplicación importante de la técnica de la invención está relacionada con el estudio y seguimiento del fenómeno de memoria de una cuasiespecie en un organismo infectado y su correlación con la dinámica poblacional de un virus.

45 Una ventaja interesante de la técnica de la invención es que permite cuantificar genomas minoritarios, especialmente genomas memoria minoritarios, presentes en una cuasiespecie viral mediante la incorporación de los controles adecuados, por ejemplo:

[1] Series por triplicado de mezclas a distintas concentraciones conocidas de oligonucleótidos de secuencia igual a la del virus salvaje y/o a la secuencia mayoritaria de la cuasiespecie y a una secuencia mutante. Una serie característica contiene un oligonucleótido del virus mutante en proporciones de 1, $5 \cdot 10^{-1}$, 10^{-1} , $5 \cdot 10^{-2}$, 10^{-2} , $5 \cdot 10^{-3}$, 10^{-3} , $5 \cdot 10^{-4}$, 10^{-4} , $5 \cdot 10^{-5}$, 10^{-5} , y 0, respecto a un oligonucleótido salvaje. Tras la hibridación con una muestra marcada del virus salvaje y posterior escaneado se obtiene un patrón de intensidades decrecientes, de manera que todos los puntos que queden en la zona lineal de la curva permiten construir una recta de calibrado y con ello establecer una relación directa entre intensidad de hibridación y proporción de moléculas respecto a un total. La intensidad de la señal de hibridación dada por los oligonucleótidos minoritarios, por ejemplo, los oligonucleótidos memoria minoritarios, con posiciones interrogantes tras la hibridación de la muestra test marcada con otro fluorocromo, se puede correlacionar con los valores obtenidos en la curva patrón;

[2] Puntos por cuadruplicado comprendidos por oligonucleótidos iguales a las mismas concentraciones, complementarios a los oligonucleótidos contenidos en diferentes concentraciones en una mezcla de marcaje. Ni estos oligonucleótidos ni sus complementarios deben formar parte de los fragmentos amplificados por PCR del virus salvaje y de la muestra test. La correlación entre la intensidad de la señal de hibridación y la obtenida para los oligonucleótidos minoritarios, por ejemplo, oligonucleótidos memoria minoritarios, hibridados con la muestra test, permite estimar la proporción de los genomas mutantes presentes en la cuasiespecie estudiada; y

65 [3] Series de oligonucleótidos con posiciones interrogantes con posiciones complementarias, pero no las secuencias flanqueantes, están ausentes en las mezclas de marcaje del virus salvaje. La media de las intensidades de hibridación dada por estas series constituye el “ruido de fondo” de la hibridación.

Una ventaja adicional de la invención es que permite identificar genomas minoritarios surgidos de fenómenos diferentes al genoma memoria pero que pueden tener implicaciones importantes en la persistencia del virus en el organismo para ser identificado.

5 La invención también proporciona un kit para la detección de genomas minoritarios presentes en cuasiespecies virales que incluye, al menos, un oligonucleótido que sirve como un control positivo, al menos, uno que sirve como un control negativo, al menos, un oligonucleótido memoria seleccionado entre los oligonucleótidos identificados como OM1, OM2, OM3, OM4, OM5, OM6, OM7 y OM8; y unos medios que permitan realizar una curva de calibrado. Los distintos oligonucleótidos presentes en el kit se encuentran, preferentemente, fijados covalentemente y de forma
10 ordenada en un microchip de ADN. La posición interrogante de los oligonucleótidos memoria se refiere a aquellas mutaciones responsables de la resistencia a fármacos o del escape al sistema de defensa del organismo infectado. También se incorporan oligonucleótidos memoria con posiciones interrogantes de mutaciones aparentemente mutaciones silenciosas que pueden conferir resistencia a fármacos frente a los cuales el virus ha sido expuesto todavía, o que pueden producir, por combinación con otras mutaciones posteriores, epítomos nuevos no reconocidos por el sistema inmunológico. Asimismo, el kit proporcionado por la invención puede incluir un juego de oligonucleótidos para amplificar por RT-PCR y/o por PCR o por PCR anidada los fragmentos de la secuencia del genoma viral donde se van a localizar las mutaciones. El kit proporcionado por la invención puede contener, además, la totalidad o parte de los reactivos necesarios para la realización del método descrito en la invención, incluyendo las soluciones tampón apropiadas, así como instrucciones, protocolos, consejos prácticos con una guía de problemas y soluciones, y un empaquetamiento
20 adecuado.

En una realización particular, el kit proporcionado por la presente invención es un kit para la detección de genomas minoritarios, preferentemente, genomas memoria minoritarios, presentes en cuasiespecies del VIH-1, VIH-2, VHC y VHB. A modo de ejemplo, los oligonucleótidos memoria que contienen las posiciones interrogantes de las mutaciones mostradas en la Tabla II (véase el Ejemplo 2) permiten identificar mutaciones de memoria asociadas a resistencia en los genes que codifican la proteasa (PR) y la transcriptasa inversa (RT) del VIH. La presente invención incluye el empleo de todos los oligonucleótidos con posiciones interrogantes en cada una de las tres posiciones de los codones de los genes de la PR y la TR del VIH que aparecen en la Tabla II. En otra realización particular, la invención incluye el empleo de todos los oligonucleótidos con posiciones interrogantes para todas las posiciones incluidas en las regiones 1978-2010 y 6625-6744 (tomando la A del ATG de la poliproteína como posición 1), y 175-350 (respecto la posición 1 del genoma viral) del VHC tipo 1-b.
30

Fabricación de microchips de ADN

35 La fabricación de microchips de ADN que contienen oligonucleótidos interrogantes para la identificación de genomas minoritarios, por ejemplo, genomas memoria minoritarios, presentes en cuasiespecies virales se realiza mediante métodos convencionales. En general, los oligonucleótidos interrogantes se pueden obtener de casas comerciales por síntesis química. El soporte sólido del microchip puede ser un portaobjetos de vidrio para microscopía en el que se deposita mediante un sistema robotizado una muestra de entre 0,2 y 30 nl de oligonucleótido conteniendo de 2 a 300 fmol y formando un punto de 50 a 250 μm de diámetro. El microchip debería estar formado por sondas que identifiquen la copia tipo (como control), sondas que identifiquen los genomas mayoritarios y minoritarios (o memoria), así como sondas que sirvan de control negativo, es decir, que no den hibridación ni con la copia maestra ni con los genomas minoritarios (o memoria). De forma similar, se incluye una curva de calibrado determinada mediante concentraciones conocidas de secuencias diana marcadas, tal como se ha indicado anteriormente. El número de oligonucleótidos presentes en el chip depende del número de posiciones interrogadas. En general, se requieren al menos 10 oligonucleótidos interrogantes por cada uno de los codones del gen interrogado [véase la Tabla I], un oligonucleótido para la copia maestra, y tres con el resto de posiciones interrogantes para cada una de las posiciones del codón. De esta forma, un microchip para buscar genomas minoritarios o memoria con mutaciones en 200 posibles posiciones debería de tener al menos 2.000 sondas. No obstante, ya que la síntesis de grandes cantidades de oligonucleótidos resulta económicamente caro, el número de puntos del microchip se puede disminuir considerablemente si lo que se buscan son genomas memoria de secuencia conocida (como los mostrados en la Tabla II), o si se hace un estudio previo en el que se eliminan las posibles mutaciones que no suponen cambio de aminoácido en la proteína sintetizada. Un diseño típico de OM1 se muestra en la Tabla I.
50

55

60

65

ES 2 273 835 T3

TABLA I

Diseño de oligonucleótidos memoria de tipo 1 (OM1)

5	Mutación	Oligonucleótido
10	a) Para una sola mutación:	
15	ATG → TTG	NNNNNNNTACNNNNN NNNNNNNAACNNNNN mut
20		
25	b) Para mutaciones en las tres posiciones:	
30	ATG → NTG, ANG, ATN	NNNNNNNTACNNNNN NNNNNNNAACNNNNN mut1 NNNNNNNCACNNNNN mut2 NNNNNNNGACNNNNN mut3 NNNNNTACNNNNNN NNNNNTCCNNNNNN mut4 NNNNNTGCNNNNNN mut5 NNNNNTTCNNNNNN mut6 NNNNNTACNNNNNN NNNNNTAANNNNNNN mut7 NNNNNTAGNNNNNNN mut8 NNNNNTATNNNNNNN mut9

Donde:

- N representa cualquiera de los cuatro nucleótidos (A,C,G,T);
- wt, es el salvaje; y
- mut, es mutante.

Existen distintos protocolos para fijar el ADN al soporte y prepararlo para la hibridación con las muestras diana marcadas. La hibridación de las muestras test y de referencia con las sondas del microchip se lleva a cabo en condiciones específicas, de tal forma que se formen híbridos estables con sondas en las que la complementariedad sea del 100%. Estas condiciones de hibridación dependen del tipo de oligonucleótidos minoritarios o memoria presentes en el microchip. Por ejemplo, para OM2, para la unión del primero de los oligonucleótidos se puede utilizar un tampón de hibridación comprendido por NaCl 1M, EDTA 1 mM, Tween 20 1% y fosfato sódico 5 mM, pH 7,0 (Parinov *et al.*, 1996), y el tiempo de hibridación es de 15 minutos a 0°C. El ADN que no se ha unido se lava con el mismo tampón de hibridación a la misma temperatura durante 10-20 segundos. La unión del segundo oligonucleótido estancado se lleva a cabo a 0°C durante 5 minutos en el mismo tampón de hibridación y los oligonucleótidos no unidos se lavan con el mismo tampón a 20°C durante 10 minutos. Las condiciones del segundo lavado son más restrictivas puesto que los híbridos formados por el apilamiento de los dos oligonucleótidos son más estables en estas condiciones.

Las condiciones de hibridación y de lavado deberían ser optimizadas para cada tipo de oligonucleótido. El kit de la presente invención lleva incorporado unos standars, protocolos y todo lo necesario para proporcionar las condiciones requeridas para la correcta utilización la correcta utilización del kit. De igual forma, el kit también incorpora los objetos necesarios para establecer las correctas condiciones para llevar a cabo las modificaciones enzimáticas para detectar mutaciones memoria con la técnica de PCR/LDR. También se contempla la posibilidad de modificaciones enzimáticas al microchip después de la hibridación, con el objetivo de mejorar la relación señal/ruido de la misma.

A continuación se detallan las referencias bibliográficas descritas en el texto:

Allen, M., Deslauriers, M., Andrews, C., et al. (1998). Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. *Hepatology* 27: 1670-1677.

Baranowski, E., Sevilla, N., Verdaguer, N., Ruiz-Jarabo, C. M., Beck, E., and Domingo, E. (1998). Multiple virulence determinants of foot-and-mouth disease virus in cell culture. *J. Virol* 72, 6362-6372.

Batschelet, E., Domingo, E., and Weissmann, C. (1976). The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. *Gene* 1, 27-32.

Blanchard, A. P., Kaiser, R. J. and Hood, L. E. (1996). Synthetic DNA arrays. *Biosensors and Bioelectronics* 11, 687-690.

Boom, R., Sol, C., Salimans, M., et al. (1990). Rapid and simple method for purification of nucleic acids, *J. Clin. Microbiol.* 28: 495-503.

Borrow, P., Lewicki, H., Wei, X., Horwitz, M. S., Pfeffer, N., Meyers, H., Nelson, J. A., Gairin, J. E., Hahn, B. H., Oldstone, M. B., and Shaw, G. M. (1997). Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat. Med.* 3, 205-211.

Borrow, P., and Shaw, G. M. (1998). Cytotoxic T-lymphocyte escape viral variants: how important are they in viral evasion of immune clearance *in vivo*? *Immunol. Rev.* 164, 37-51.

Briones, C., Mas, A., Gómez-Mariano, G., Altisent, C., Menéndez-Arias, L., Soriano, V., and Domingo, E. (2000). Dynamics of dominance of a dipeptide insertion in reverse transcriptase of HIV-1 from patients subjected to prolonged therapy. *Virus Res.* 66, 13-26.

Bulyk, M. L., Gentalen, E. Lockhart, D. J. and Church, M. 1999. DNA-protein interactions by double-stranded DNA arrays. *Nature Biotech.* 17, 573-577.

Cabot, B., Martell, M., Esteban, J.I., Sauleda, S., Otero, T., Esteban, R., Guardia, J. y Gómez, J., 2000. Nucleotide and aminoacid complexity of hepatitis C virus quasispecies in serum and liver. *J. Virol.* 74, 805-811.

Chen Z, Luckay A, Sodora DL, et al (1997). Humman immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys, *J. Virol.* 7: 3953-3960.

Chen, W., and Baric, R. S. (1996). Molecular anatomy of mouse hepatitis virus persistence: coevolution of increased host cell resistance and virus virulence. *J. Virol* 70, 3947-3960.

Clavel F, Guétard D, Brun-Vézinet F, et al (1986). Isolation of a new human retrovirus from West African patients with AIDS, *Science* 233: 343-346.

Coleman, P., McQuillan, G., Moyer, L., et al. (1998). Incidence of hepatitis B virus infection in the United States, 1976-1994: estimates from the National Health and Nutrition Examination Surveys. *J. Infect. Dis.* 178: 954-959.

- Cornelissen, M., van den Burg, R., Zorgdrager, F., Lukashov, V., and Goudsmit, J. (1997).** pol gene diversity of five human immunodeficiency virus type 1 subtypes: evidence for naturally occurring mutations that contribute to drug resistance, limited recombination patterns, and common ancestry for subtypes B and D. *J. Virol* 71, 6348-6358.
- 5 **De Antoni A, Foli A, Lisiewicz J, et al (1997).** Mutations in the pol gene of human immunodeficiency virus type 1 in infected patients receiving didanosine and hydroxyurea combination therapy, *J. Infect. Dis.* 176: 899-903.
- De Jong JJ, Goudsmit J, Lukashov VV, et al (1999).** Insertion of two aminoacids combined with changes in reverse transcriptase containing tyrosineB215 of HIV-1 resistant to multiple nucleoside analogs, *AIDS* 13: 75-80.
- 10 **De Watcher, R., and Fiers, W. (1972).** Preparative two-dimensional polyacrylamide gel electrophoresis of 32P-labeled RNA. *Anal. Biochem.* 49, 184-197.
- Domingo, E., Flavell, R. A., and Weissmann, C. (1976).** *In vitro* site-directed mutagenesis: generation and properties of an infectious extracistronic mutant of bacteriophage Q β . *Gene* 1, 3-25.
- 15 **Domingo, E., Sabo, D., Taniguchi, T., and Weissmann, C. (1978).** Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13, 735-744.
- Domingo, E., Díez, J., Martínez, M. A., Hernández, J., Holguín, A., Borrego, B., and Mateu, M. G. (1993).** New observations on antigenic diversification of RNA viruses. Antigenic variation is not dependent on immune selection [published erratum appears in *J Gen Virol* 1994 Apr;75(Pt 4):949]. *J. Gen. Virol* 74, 2039-2045.
- 20 **Domingo, E. (1996).** Biological significance of viral quasispecies. *Viral Hepatitis Reviews* 2, 247-261.
- 25 **Domingo, E. (1997b).** RNA virus evolution, population dynamics, and nutritional status. *Biol. Trace Elem. Res* 56, 23-30.
- Domingo, E. (1999a).** Quasispecies. In “Encyclopedia of Virology” (A. Granoff, and R. G. Webster, Eds.), pp. 1431-1436. *Academic Press*, London.
- 30 **Domingo, E., Webster, R. G., and Holland, J. J., Eds. (1999d).** Origin and Evolution of Viruses. San Diego: *Academic Press*.
- Domingo, E., Biebricher, C., Holland, J. J., and Eigen, M. (2000).** Quasispecies and RNA Virus Evolution: Principles and Consequences. Landes Bioscience, Austin. *Landes Bioscience*, Austin.
- 35 **Eigen, M. (1971).** Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465-523.
- 40 **Eigen, M. (1992).** Steps towards life. *Oxford University Press*.
- Eigen, M. (1996).** On the nature of virus quasispecies. *Trends. Microbiol.* 4, 216-218.
- 45 **Eigen, M., and Schuster, P. (1977).** The hypercycle: a principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften* 64, 541-565.
- Eigen, M., and Schuster, P. (1978a).** The hypercycle: a principle of natural self-organization. Part B: The abstract hypercycle. *Naturwissenschaften* 65, 7-41.
- 50 **Eigen, M., and Schuster, P. (1978b).** The hypercycle: a principle of natural self-organization. Part C: The realistic hypercycle. *Naturwissenschaften* 65, 341-369.
- Eigen, M., and Schuster, P. (1979).** The hypercycle. A principle of natural self-organization. *Springer*.
- 55 **Eigen, M., McCaskill, J., and Schuster, P. (1988).** Molecular quasi-species. *J. Phys. Chem.* 92, 6881-6891.
- Escarmís, C., Dávila, M., and Domingo, E. (1999).** Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller’s ratchet. *J. Mol. Biol* 285, 495-505.
- 60 **Escarmís, C., Dávila, M., Charpentier, N., Bracho, A., Moya, A., and Domingo, E. (1996).** Genetic lesions associated with Muller’s ratchet in an RNA virus. *J. Mol. Biol.* 264, 255-267.
- Evans, D. T., O’Connor, D. H., Jing, P., Dzuris, J. L., Sidney, J., da Silva, J., Allen, T. M., Horton, H., Venham, J. E., Rudersdorf, R. A., Vogel, T., Pauza, C. D., Bontrop, R. E., DeMars, R., Sette, A., Hughes, A. L., and Watkins, D. I. (1999).** Virus-specific cytotoxic T-lymphocyte responses select for amino-acid variation in simian immunodeficiency virus Env and Nef. *Nature Medicine* 5, 1270-6.
- 65

- Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B., and Straus, S. E., Eds. (1996).** Field's Virology 3rd edition. Philadelphia: Lippincott-Raven.
- Flint, S. J., Enquist, L. W., Krug, R. M., Racaniello, V. R., and Skalka, A. M. (2000).** Virology. Molecular Biology, Pathogenesis and Control. *ASM Press*, Washington D.C.
- Forns, X., Purcell, R. H., and Bukh, J. (1999).** Quasispecies in viral persistence and pathogenesis of hepatitis C virus. *Trends Microbiol.* 7, 402-410.
- Gao F. et al (1998).** A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1, *J. Virol* 72: 5680-5698.
- Gerotto M, Sullivan DG, Polyak SJ, et al (1999).** Effect of Retreatment with interferon alone or interferon plus rivabirin on Hepatitis C Virus quasispecies diversification in nonresponder patients with chronic hepatitis C, *J. Virol.* 73: 7241-7247].
- Gerry, N.P., Witowski, N.E., Day, J., Hammer, R.P., Barany, G. and Barany, F. (1999).** Universal DNA microarray method for multiplex detection of low abundance point mutations. *J. Mol. Biol.* 292, 251-262.
- Gibbs, A., Calisher, C., and García-Arenal, F., Eds. (1995).** Molecular Basis of Virus Evolution. Cambridge: Cambridge University Press.
- Granoff, A., and Webster, R. G., Eds. (1999).** Encyclopedia of Virology 2nd edition. San Diego: Academic Press.
- Hacia J.G., Sun B., Hunt N., et al (1999).** Strategies for mutational analysis of the large multi-exon ATM gene using high-density oligonucleotide arrays, *Genome Research* 8: 1245-1258.
- Hacia, J.G. (1999).** Resequencing and mutational analysis using oligonucleotide microarrays. *Nature Genetics* 21, 42-47.
- Havlir, D. V., Eastman, S., Gamst, A., and Richman, D. D. (1996).** Nevirapine-resistant human immunodeficiency virus: kinetics of replication and estimated prevalence in untreated patients. *J. Virol.* 70, 7894-7899.
- Heitman, J. (1993).** On the origins, structures and functions of restriction-modification enzymes. In "Genetic Engineering" (J. K. Setlow, Ed.), Vol. 15, pp. 57-108. *Plenum Press*, New York.
- Hertogs K, Bloor S, de Vroey V, et al (2000).** A novel human immunodeficiency virus type I reverse transcriptase mutational pattern confers phenotypic lamivudine resistance in the absence of mutation 184V, *Antimicrob. Agents Chemother.* 44: 568-573.
- Howe, C. J., and Ward, E. S. (1989).** Nucleic Acids Sequencing: A Practical Approach. *IRL Press*, Oxford.
- Innis, M.A., Gefland, D.H., Sninsky, J.J. y White, T.J.** PCR protocols. A guide to methods and applications. *Academic Press*, San Diego, 1990.
- Janssens W, Heyndrickx L, Van der Auwera G, et al (1999).** Interpatient genetic variability of HIV-1 group O, *AIDS* 13: 41-48.
- Korber B, Kuiken C, Foley B, Hahn B, McCutchan F, Mellors JW y Sodroski J,** Human Retroviruses and AIDS 1998, Los Alamos National Laboratory, Los Alamos, New Mexico, USA, 1999; actualizaciones periódicas en <http://hiv-web.lanl.gov>.
- Liaw, Y. F., Tai, D. I., Chu, C. M., et al. (1988).** The development of cirrhosis in patients with chronic type B hepatitis: a prospective study. *Hepatology* 8: 493-496.
- Luria, S. E., Darnell, J., J.E., Baltimore, D., and Campbell, A. (1978).** General Virology, 2nd ed. *John Wiley and Sons*, New York.
- Martínez, M. A., Verdaguer, N., Mateu, M. G., and Domingo, E. (1997).** Evolution subverting essentiality: dispensability of the cell attachment Arg-Gly-Asp motif in multiply passaged foot-and-mouth disease virus. *Proc Natl Acad Sci USA* 94, 6798-6802.
- Mas A, Quiñones-Mateu ME, Domingo E y Soriano V (1999).** Phylogeny of HIV type 1 group O isolates based on env gene sequences, *AIDS Res. Hum. Retrov.* 15: 769-773.
- Maskos, U. and Southern E. M. 1993.** A novel method for the analysis of multiple sequence variants by hybridisation to oligonucleotide arrays. *Nucleic Acid Res.* 21, 2267-2268.

- Mateu, M. G., Martínez, M. A., Rocha, E., Andreu, D., Parejo, J., Giralt, E., Sobrino, F., and Domingo, E.** (1989). Implications of a quasispecies genome structure: effect of frequent, naturally occurring amino acid substitutions on the antigenicity of foot-and-mouth disease virus. *Proc Natl Acad Sci USA* 86, 5883-5887.
- 5 **Mateu, M. G., Rocha, E., Vicente, O., Vayreda, F., Navalpotro, C., Andreu, D., Pedroso, E., Giralt, E., Enjuanes, L., and Domingo, E.** (1987). Reactivity with monoclonal antibodies of viruses from an episode of foot-and-mouth disease. *Virus Res* 8, 261-274.
- Maynard, J.E.** (1990). Hepatitis B: global importance and need for control. *Vaccine* 8: S18-S20.
- 10 **McCutchan F, Salminen M, Carr J y Burke D** (1996). HIV-1 genetic diversity, *AIDS* 10: S13-S20
- McMichael, A. J., and Phillips, R. E.** (1997). Escape of human immunodeficiency virus from immune control. *Annu. Rev. Immunol.* 15, 271-296.
- 15 **Menéndez-Arias, L., and Domingo, E.** (1998). Resistance tables for antiretroviral drugs, *AIDS Cyber J.* 1: 95-127.
- Menéndez-Arias, L., and Domingo, E.** (1999). Cambios de aminoácido asociados a resistencia a inhibidores de la retrotranscriptasa y la proteasa del VIH. In "Guía práctica para el manejo clínico de las resistencias al VIH" (B. Clotet, L. Ruiz, X. Martínez-Picado, L. Menéndez-Arias, J. M. Gatell, and D. D. Richman, Eds.), pp. 215-267. Grupo E, Enters, Merck Sharp and Dohme de España, S.A.
- Mohanty, S.B., and Dutta, S.K.** (1981). *Veterinary Virology. Lea & Febiger, Philadelphia.*
- 25 **Morse, S. S., Ed.** (1993). *Emerging viruses: Oxford University Press, Oxford.*
- Morse, S. S., Ed.** (1994). *The Evolutionary Biology of Viruses. New York: Raven Press.*
- 30 **Mortara, L., Letourneur, F., Gras-Masse, H., Venet, A., Guillet, J. G., and Bourgault-Villada, I.** (1998). Selection of virus variants and emergence of virus escape mutants after immunization with an epitope vaccine. *J Virol* 72, 1403-1410.
- Murphy, F. A.** (1996). Virus taxonomy. In "Field's Virology" (B. N. Fields, D. M. Knipe, P. M. Howley, R. M. Chanock, J. L. Melnick, T. P. Monath, B. Roizman, and S. E. Straus, Eds.), pp. 15-59. *Lippincott-Raven Publishers, Philadelphia Pa.*
- 35 **Nájera, I., Holguín, A., Quiñones-Mateu, M. E., Muñoz-Fernández, M. A., Nájera, R., López-Galíndez, C., and Domingo, E.** (1995). Pol gene quasispecies of human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy. *J. Virol.* 69, 23-31.
- Palmer, S., Shafer, R. W., and Merigan, T. C.** (1999). Highly drug-resistant HIV-1 clinical isolates are cross-resistant to many antiretroviral compounds in current clinical development. *AIDS* 13, 661-7.
- 45 **Paraskevis D y Hatzakis A** (1999). Molecular epidemiology of HIV-1 infection, *AIDS Rev.* 1: 238-249.
- Parinov, S., Barsky, V., Yershov, G., Kirillov, E., Timofeev, E., Belgovskiy, A. and Mirzabekov, A.** 1996. DNA sequencing by hybridisation to microchip acto- and decanucleotides extended by stacked pentanucleotides. *Nucleic Acids Res.* 24, 2998-3004.
- 50 **Pawlotsky, J. M., Germanidis, G., Neumann, A. U., Pellerin, M., Frainais, P. O., and Dhumeaux, D.** (1998). Interferon resistance of hepatitis C virus genotype 1b: relationship to nonstructural 5A gene quasispecies mutations. *J. Virol.* 72, 2795-2805.
- 55 **Quiñones-Mateu, M. E., Holguín, A., Dopazo, J., Nájera, I., and Domingo, E.** (1996a). Point mutant frequencies in the pol gene of human immunodeficiency virus type 1 are two- to threefold lower than those of env. *AIDS Res Hum Retroviruses* 12, 1117-1128.
- Quiñones-Mateu, M. E., Holguín, A., Soriano, V., and Domingo, E.** (1996b). env gene diversity of HIV type 1 isolates from Spain. *AIDS Res Hum Retroviruses* 12, 955-957.
- 60 **Ratner L, Haseltine W, Patarca R, et al.** (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III, *Nature* 313: 277-284.
- 65 **Richman, D. D.** (1994). Resistance, drug failure, and disease progression. *AIDS Res. Hum. Retrov.* 10, 901-5.
- Robart, H. A.** (1995). *Human Enterovirus infections. ASM Press, Washington, D.C.*

Robertson D, Sharp P, McCutchan y Hahn B (1995). Recombination in HIV-1, *Nature* 374: 124-126.

Ruiz-Jarabo, C. M., Sevilla, N., Davila, M., Gomez-Mariano, G., Baranowski, E., and Domingo, E. (1999). Antigenic properties and population stability of a foot-and-mouth disease virus with an altered Arg-Gly-Asp receptor-recognition motif. *J. Gen. Virol.* 80, 1899-909.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning. A Laboratory Manual. *Cold Spring Harbor Laboratory Press*, New York.

Schinazi RF, Larder BA y Mellors JW (1999). Mutations in retroviral genes associated with drug resistance: 1999-2000 update, *Internatl. Antiviral News* 7: 46-69.

Schinazi, R. F., Larder, B. A., and Mellors, J. W. (1997). Mutations in retroviral genes associated with drug resistance. *Int. Antivir. News* 5, 129-142.

Schuster, P., and Stadler, P. F. (1999). Nature and evolution of early replicons. In "Origin and Evolution of Viruses" (E. Domingo, R. G. Webster, and J. J. Holland, Eds.), pp. 1-24. *Academic Press*, San Diego.

Simon F, Mauclore P, Roques P, et al, (1998). Identification of a new human immunodeficiency virus type 1 distinct from group M and group O, *Nat. Med.* 4: 1032-1037.

Southern E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol.* 98(3):503-17.

Southern et al. 1994. Arrays for complementary oligonucleotides for analysing the hybridization behavior of nucleic acids. *Nucleic Acids Res.* 22, 1368-1373, 1994.

Stocker, M., y Macpherson, I. 1964. *Nature.* 203, 1355-1357.

Taboga, O., Tami, C., Carrillo, E., Núñez, J. I., Rodríguez, A., Saíz, J. C., Blanco, E., Valero, M. L., Roig, X., Camarero, J. A., Andreu, D., Mateu, M. G., Giralt, E., Domingo, E., Sobrino, F., and Palma, E. L. (1997). A large-scale evaluation of peptide vaccines against foot-and-mouth disease: lack of solid protection in cattle and isolation of escape mutants. *J. Virol.* 71, 2606-2614.

Takehisa J, Zekeng L, Ido E, et al (1999). Human immunodeficiency virus type 1 intergroup (M/O) recombination in Cameroon, *J. Virol.* 73: 6810-6820.

Weidt, G., Deppert, W., Utermohlen, O., Heukeshoven, J., and Lehmann-Grube, F. (1995). Emergence of virus escape mutants after immunization with epitope vaccine. *J. Virol.* 69, 7147-7151.

Weiner, A., Erickson, A. L., Kansopon, J., Crawford, K., Muchmore, E., Hughes, A. L., Houghton, M., and Walker, C. M. (1995). Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escape variant. *Proc. Natl. Acad. Sci. USA* 92, 2755-2759.

Winters MA, Coolley KL, Girard YA, et al (1998). A 6-basepair insert in the reverse transcriptase gene of human immunodeficiency virus type 1 confers resistance to multiple nucleoside inhibitors, *J. Clin. Invest.* 102: 1769-1775.

Descripción de las figuras de la invención

Figura 1.- A) Descripción de las señales de hibridación obtenidas mediante un microchip de ADN con los genomas VFA-RGD y VFA-RED para distintas concentraciones de los oligonucleótidos VFA-WT-11 y VFA-MUT-11. T.hib: Temperatura de hibridación; T.lav: Temperatura de lavado. Cifras superiores: 2,5, 5, 12,5, 25, 37,5 y 50 representan las concentraciones (micromolar) de los oligos depositados. B) Cuantificación de las señales de hibridación obtenidas en los distintos puntos para cada uno de los genomas VFA-RGD y VFA-RED. Las señales de hibridación se han analizado mediante el empleo del Scanner modelo GMS 418 Array Scanner de la marca Genetic MicroSystems. La cuantificación de las señales (se considera siempre el valor medio de los dos puntos de cada duplicado) se ha realizado por medio del paquete "Statistics" del software "Array Scanner" suministrado con el equipo.

Figura 2.- A) Descripción de las señales de hibridación obtenidas mediante un microchip de ADN con los genomas VFA-RGD y VFA-RED, en una serie de puntos (duplicados) que contienen mezclas de los oligonucleótidos VFA-WT-11 y VFA-MUT-11 en distintas proporciones relativas (entre el 100% y el 0%). Las cifras superiores indican en cada caso el porcentaje del oligo VFA-WT-11 o VFA-MUT-11 en la mezcla. B) Curva de cuantificación de las señales de hibridación obtenidas. Las señales de hibridación se han analizado y cuantificado de la manera indicada en la descripción de la figura 1B.

Figura 3.- A) Descripción de las señales de hibridación obtenidas mediante un microchip de ADN con los genomas VIH-T215 y VIH-Y215 para distintas concentraciones de los oligonucleótidos VIH-WT-12 y VIH-MUT-12. T.hib: Temperatura de hibridación; T.lav: Temperatura de lavado. Cifras superiores: 2,5, 5, 12,5, 25, 37,5 y 50 representan las

ES 2 273 835 T3

concentraciones (micromolar) de los oligos depositados. B) Cuantificación de las señales de hibridación obtenidas en los distintos puntos para cada uno de los genomas VIH-T215 y VIH-Y215. Las señales de hibridación se han analizado y cuantificado de la manera indicada en la descripción de la figura 1B.

5 Figura 4.- A) Descripción de las señales de hibridación obtenidas mediante un microchip de ADN con los genomas VIH-T215 y VIH-Y215, en una serie de puntos (duplicados) que contienen mezclas de los oligonucleótidos VIH-WT-12 y VIH-MUT-12 en distintas proporciones relativas (entre el 100% y el 0%). Las cifras superiores indican en cada caso el porcentaje del oligo VIH-WT-12 o VIH-MUT-12 en la mezcla. B) Curva de cuantificación de las señales de hibridación obtenidas. Las señales de hibridación se han analizado y cuantificado de la manera indicada en la descripción de la figura 1B.

15 Figura 5.- Descripción de las secuencias de VHC implicadas en respuesta a interferón y con actividad ribozima. A) Alineamiento de secuencias de ribozima de genotipos 1-a y 1-b del VHC implicadas en respuesta a interferón. B) Secuencias del VHC relacionadas con respuesta a INF y actividad ribozima.

Ejemplos de realización de la invención

20 Los siguientes ejemplos sirven para ilustrar la invención y no deben ser considerados como limitativos del alcance de la misma.

Ejemplo 1

Detección y caracterización de genomas minoritarios del virus de la fiebre aftosa (VFA)

25 Como prueba de la presente invención original, se han construido microchips de ADN que contienen oligonucleótidos específicos de la región comprendida entre los aminoácidos 138 a 148 de la proteína VPI de la cápsida vírica (posiciones nucleotídicas 3609 a 3651 del genoma de VFA).

30 Las características de los microchips construidos son las siguientes:

35 - Se han sintetizado 4 oligonucleótidos con longitud de 11 o 15 nucleótidos (nt), todos ellos con la posición interrogante en el nucleótido central y las secuencias flanqueantes idénticas al genoma vírico original. Los oligonucleótidos han sido unidos químicamente a una amina primaria ("C6 aminolinker", también denominada "C6") en su extremo 5', con objeto de que reaccione con los grupos aldehído libres resultantes del tratamiento previo del vidrio. Entre la amina primaria y la secuencia de 11 o 15 nt existe un espaciador de 15 timidinas (T₁₅) para facilitar la hibridación. La secuencia de los cuatro oligonucleótidos utilizados se indican a continuación:

40 - "VFA-WT-15": 5'-C6-T₁₅CAAATCCCCGCGTGC-3'

45 - "VFA-MUT-15": 5'-C6-T₁₅CAAATCCTCGCGTGC-3'

50 - "VFA-WT-11": 5'-C6-T₁₅AATCCCCGCGT-3'

55 - "VFA-MUT-11": 5'-C6-T₁₅AATCCTCGCGT-3'

60 - Los oligonucleótidos presintetizados se inmovilizan en posiciones predeterminadas de los portaobjetos de vidrio mediante un sistema robotizado ("DNA arrayer").

65 - Cada uno de los cuatro oligonucleótidos ha sido depositado por duplicado en seis puntos del microchip, a concentraciones finales de 2,5, 5, 12,5, 25, 37,5 y 50 μM (micromolar) respectivamente.

70 - Además, se depositan también puntos por duplicado de las mezclas VFA-WT-15/VFA-MUT-15 y VFA-WT-11/VFA-MUT-11 en las proporciones siguientes: 100:0, 99,9:0,1, 99:1, 95:5, 90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:80, 10:90, 5:95, 1:99, 0,1:99,9, 0:100. Con ello se logra:

75 - Determinar el % de cantidad mínima detectable del genoma minoritario

80 - Realizar curvas de cuantificación

85 - Dos oligonucleótidos de 31 nt idénticos al genoma del virus salvaje ("VFA-RGD") o mutante ("VFA-RED"), marcados con el fluorocromo "Cy3" en su extremo 5'. Su secuencia se indica a continuación:

90 - "VFA-RGD": 5'-Cy3-CCGCCAGTGCACGCGGGGATTTGGCTCACCT-3'

95 - "VFA-RED": 5'-Cy3-CCGCCAGTGCACGCGAGGATTTGGCTCACCT-3'

100 Los resultados se describen en las figuras 1 y 2. En la figura 1 queda patente que el empleo de microchips de ADN permite la hibridación del oligonucleótido específico a partir de una concentración 5 μM.

ES 2 273 835 T3

La figura 2 muestra una curva de cuantificación de las señales de hibridación obtenidas en los distintos puntos con mezclas de oligonucleótidos, para cada uno de los genomas VFA-RGD y VFA-RED. A partir del punto de inflexión de dichas curvas (que corresponde a un valor de señal de en torno a 300 en ambos casos y que indica la aparición de una señal específica por encima del fondo) queda claro que el oligonucleótido minoritario en la mezcla puede detectarse incluso cuando éste sólo corresponde del 1 al 5% de la mezcla (hibridación de VFA-RGD) o del 10-20% (hibridación de VFA-RED).

Ejemplo 2

10 *Detección y caracterización de genomas minoritarios del virus de la inmunodeficiencia humana (VIH) portadores de mutaciones de resistencia a zidovudina (AZT)*

Como prueba de la presente invención, se han construido microchips de ADN que contienen oligonucleótidos específicos de VIH para detectar la mutación T215Y de resistencia a AZT. Tales oligonucleótidos son complementarios a la región comprendida entre los aminoácidos 210 a 220 de la transcriptasa inversa de VIH-1 (posiciones nucleotídicas 15 3179 a 3211 del aislado CAM-1 de VIH-1).

Las características de los microchips construidos son las siguientes:

20 - Se han sintetizado 4 oligonucleótidos con longitud de 12 o 16 nucleótidos (nt), todos ellos con las dos posiciones interrogantes contiguas (ACC→TAC) en su centro, y las secuencias flanqueantes idénticas al genoma vírico. Los oligonucleótidos han sido unidos químicamente a una amina primaria ("C6 aminolinker", también denominada "C6") en su extremo 5', con objeto de que pueda reaccionar con los grupos aldehído libres resultantes del tratamiento previo del vidrio. Entre la amina primaria y la secuencia de 12 o 16 nt existe un espaciador de 15 timidinas (T₁₅) para facilitar la hibridación. La secuencia de los cuatro oligonucleótidos utilizados se indican a continuación:

- "VIH-WT-16": 5'-C6-T₁₅TGGTGTGGTAAGTCCC-3'

- "VIH-MUT-16": 5'-C6-T₁₅TGGTGTGTAAAGTCCC-3'

30 - "VIH-WT-12": 5'-C6-T₁₅GTGTGGTAAGTC-3'

- "VIH-MUT-12": 5'-C6-T₁₅GTGTGTAAAGTC-3'

35 - Los oligonucleótidos presintetizados se inmovilizan en posiciones predeterminadas de los portas de vidrio mediante un sistema robotizado ("DNA arrayer").

- Cada uno de los cuatro oligonucleótidos ha sido depositado por duplicado en seis puntos del microchip, a concentraciones finales de 2,5, 5, 12,5, 25, 37,5 y 50 μM respectivamente.

40 - Además, se depositan también puntos por duplicado de las mezclas VIH-WT-16/VIH-MUT-16 y VIH-WT-12/VIH-MUT-12 en las proporciones siguientes: 100:0, 99,9:0,1, 99:1, 95:5, 90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:80, 10:90, 5:95, 1:99, 0,1:99,9, 0:100. Con ello se logra:

45 - Determinar el % mínimo detectable del genoma minoritario

- Realizar curvas de cuantificación

50 - Se han utilizado dos oligonucleótidos de 32 nt idénticos al genoma del virus salvaje ("VIH-T215") o mutante ("VIH-Y215"), marcados con el fluorocromo "Cy3" en su extremo 5'. Su secuencia se indica a continuación:

- "VIH-T215": 5'-Cy3-TTGAGGTGGGGACTTACCACACCAGACAAAAA-3'

- "VIH-Y215": 5'-Cy3-TTGAGGTGGGGACTTTACACACCAGACAAAAA-3'

55 Los resultados se describen en las figuras 3 y 4. En la figura 3 se pone de manifiesto que, en este caso, el empleo de microchips de ADN permite la hibridación del oligonucleótido específico a partir de una concentración 2,5 μM. Asimismo, se demuestra que el grado de especificidad alcanzado es muy elevado y que las hibridaciones inespecíficas quedan con una señal que no es distinguible del fondo (siempre menor de 200 unidades), frente a las 800 o más unidades obtenidas por hibridación específica a oligos con concentración superior a 2,5 μM.

65 La figura 4 muestra una curva de cuantificación de las señales de hibridación obtenidas en los distintos puntos con mezclas de oligos, para cada uno de los genomas VIH-T215 y VIH-Y215. El punto de inflexión de las curvas (que corresponde a un valor señal de en torno a 300 en ambos casos y que indica la aparición de una señal específica por encima del fondo) demuestra la posibilidad de detectar el oligonucleótido minoritario en la mezcla incluso cuando éste está en una proporción comprendida entre el 1 y el 5% para ambos casos (hibridación de VIH-T215 Y VIH-Y215).

Ejemplo 3

Detección y caracterización de genomas memoria en poblaciones del virus de la fiebre aftosa (VFA)

5 Ejemplo 3.1

Método de detección para el genoma memoria "RED" del VFA resistente a un anticuerpo monoclonal

Se parte de un mutante del VFA que contiene la secuencia Arg-Glu-Asp (arginina-glutámico-aspartico, en adelante RED) en las posiciones 141-143 de la proteína VP1 de la cápsida vírica. Este mutante se obtuvo por aislamiento de mutantes resistentes a neutralización (pérdida de infectividad) de una población de VFA que contiene la secuencia salvaje Arg-Gly-Asp (arginina-glicina-aspartico, en adelante RGD) en las posiciones 141-143 de la proteína VP1 de la cápsida vírica (posiciones 3628-3636 del genoma del virus, según la numeración descrita en Escarmís *et al.*, 1999). Tanto el mutante con RED (en adelante VFA RED) como el virus parental RGD (en adelante VFA RGD) así como el anticuerpo monoclonal SD6 (en adelante AM SD6) empleado en la selección de VFA RED a partir de VFA RGD han sido descritos previamente (Martínez *et al.*, 1997; Mateu *et al.*, 1987; Mateu *et al.*, 1989; Ruiz-Jarabo *et al.*, 1999). Tras propagar por triplicado muestras de VFA RED en células BHK-21 (línea celular establecida a partir de células de riñón de hamster, descrita en Stocker y Macpherson, 1964) se observó que la secuencia RED revertía a RGD a medida que se multiplicaba el virus. Dado que VFA RED fue inicialmente un clon aislado de una placa viral (proveniente de un genoma único), la transición de VFA RED a VFA RGD tuvo que ser resultado de una reversión verdadera, es decir del cambio de nucleótido A425→G (el número indica la posición en la zona codificante de VP1, que es equivalente a la posición 3632 del genoma completo del VFA, según la numeración descrita en Escarmís *et al.*, 1999). La propagación del VFA RED por triplicado consistió en la infección de 4×10^6 de células BHK-21 en 4×10^5 unidades formadoras de placa infecciosas (en adelante UFP) del VFA RED. El proceso de infección seriada se repitió 25 veces (o 25 pases); en cada pase se infectaron 4×10^6 células BHK-21 con 4×10^5 UFP del virus obtenido en la infección (pase) anterior. La proporción de VFA RGD y VFA RED en los pases fue determinada mediante secuenciación de los genomas presentes en la población, en la zona del RED codificante de los aminoácidos 141 a 143 de la proteína VP1 y la zona circundante. Los métodos empleados para este análisis han sido descritos anteriormente (Escarmís *et al.*, 1996; Escarmís *et al.*, 1999; Baranowski *et al.*, 1998). Tras diez pases en las condiciones mencionadas, no se detectó la presencia de secuencias genómicas codificantes de RED en la secuencia consenso o promedio (la que se obtiene al secuenciar la población de genomas presentes en la muestra analizada, sin usar ningún proceso de clonaje molecular previo) de la población. Se quiso determinar si en los pases 15 y 25 la población mantenía en su espectro de mutantes una memoria molecular de su origen como VFA RED. Para ello se determinó la frecuencia de mutantes resistentes al AM SD6 y la secuencia de varios de estos mutantes en la zona que codifica las posiciones 141 a 143 de VP1. Para las tres poblaciones del pase 15 las frecuencias fueron: $1,8 \times 10^{-2}$, $1,4 \times 10^{-2}$ y $1,3 \times 10^{-2}$, respectivamente; y para las tres poblaciones del pase 25 las frecuencias fueron $1,6 \times 10^{-2}$, $5,3 \times 10^{-1}$ y $2,7 \times 10^{-4}$, respectivamente. Para dos poblaciones control (el mismo VFA parental y otra población clonal de la misma estirpe) las frecuencias halladas fueron $(4,1\Gamma 0,5) \times 10^{-3}$ y $(5,0\Gamma 1,6) \times 10^{-4}$, respectivamente. La prueba más definitiva de la presencia de memoria se obtuvo mediante secuenciación de 15 clones de la población revertiente del pase 15 y 15 clones de la población revertiente del pase 25, resistentes a AM SD6. Todos los mutantes analizados (30 de los 30 analizados) tenían la secuencia de nucleótidos genómica codificante de RED, mientras que en poblaciones control muy pocos mutantes resistentes al anticuerpo monoclonal SD6 incluían RED (tan sólo 4 de 112 clones analizados) ($P < 0,001$; prueba χ^2). Por tanto, la población de VFA RGD revertiente mantenía una memoria estable de la dominancia anterior del virus con RED en la historia del virus. Como prueba adicional de que la memoria de las cuasiespecies está representada en componentes minoritarios del espectro de mutantes, las poblaciones VFA RGD revertientes del pase 15 fueron pasadas en células BHK21 10 veces más, empleando entre 10 y 100 UFPs para infectar 10^6 células BHK-21 por pase resultando en la pérdida de los genomas memoria. Al secuenciar 15 mutantes resistentes al anticuerpo monoclonal, ninguno de ellos mostró la secuencia RED, sino que aparecieron sustituciones de aminoácido en otras posiciones del entorno (aminoácidos 139, 142, 143, 144 y 146). Este resultado demuestra que la memoria de las cuasiespecies es una propiedad del conjunto del espectro de mutantes y no de genomas individuales que componen la cuasiespecie.

El diseño de microchips de ADN con oligonucleótidos específicos de la región comprendida entre los aminoácidos 138 a 148 de la proteína VP1 de la cápsida vírica (posiciones 3609-3651 del genoma de VFA) permite detectar la presencia del genoma memoria minoritario RED en una cuasiespecie viral dominada por el genoma RGD, y cuantificar la proporción en que está presente en distintas condiciones experimentales. Asimismo, esto también permite detectar otros cambios nucleotídicos localizados en las regiones flanqueantes. Para ello, se diseña una colección de nucleótidos de 20 nt que cada uno posee, como posición interrogante central, cada uno de los cuatro nucleótidos (A, C, G, T) posibles en cada una de las posiciones comprendidas entre las posiciones 3609-3651 del genoma de VFA, según los métodos indicados en la descripción detallada de la invención. Las regiones flanqueantes corresponden, en todos los casos, a la secuencia del genoma salvaje (RGD) en esa región.

Ejemplo 3.2

65 *Determinación de genomas memoria del virus VFA con colas de poli-A de longitud variable entre las posiciones 1119-1123 del genoma viral*

Una prueba adicional de la presencia de memoria en cuasiespecies de VFA se obtuvo empleando el clon C₂₂⁹, altamente debilitado por pases sucesivos placa a placa del VFA, tal como se describe en (Escarmís *et al.*, 1996). Este

clon tiene como característica única, no encontrada en ningún otro aislado natural o de laboratorio de VFA, la presencia de un tramo de poliadenilato (o poli A, tramo de AMP polimerizado) de longitud heterogénea, con un promedio de unos 23 residuos de ácido adenílico (abreviadamente A). Este poli A se sitúa en la región del genoma que precede al segundo triplete de AUG que funciona en la iniciación de la síntesis de proteínas (la poliproteína) codificada por el genoma del VFA (Escarmís *et al.*, 1996). Cuando el clon C₂₂⁹ se propagó en cultivos celulares empleando elevados tamaños poblacionales (en cada pase se infectaron 4 x 10⁶ células BHK-21 con 10⁶ a 10⁷ UFPs del virus obtenido en el pase anterior) tuvo lugar un aumento de la eficacia replicativa del virus (Escarmís *et al.*, 1999). En este proceso de ganancia de eficacia replicativa, la primera alteración molecular observada fue la pérdida del poliadenilato que resultó no detectable en el pase 20 (Escarmís *et al.*, 1999). Cuando la población de clon C₂₂⁹, pasada 50 veces en células BHK-21, se sometió a clonaje molecular y biológico, se detectaron genomas con un número mayor de residuos A que el presente en el VFA salvaje, precediendo el segundo triplete AUG funcional. Se detectó un mayor número de As en 8 de los 70 clones analizados, mientras que no se detectó ningún genoma con un número mayor de As en 40 clones de una población control de VFA sometida al mismo número de pases, pero que se originó en un clon sin As adicionales (0,01 > P > 0,0025; prueba Chi²). Es decir, el VFA C₂₂⁹ mantuvo una memoria de su historia anterior en forma de genomas minoritarios de la cuasiespecie. Estos ejemplos con VFA prueban la existencia de una memoria molecular en poblaciones de VFA, que revela la historia evolutiva anterior del virus.

Ejemplo 4

20 *Método para la detección de genomas memoria del virus de la inmunodeficiencia humana (VIH) portadores de mutaciones que confieren resistencia a fármacos en enfermos tratados*

Los chips de ADN específicos para detectar genomas memoria del Virus de la Inmunodeficiencia Humana (VIH) contienen una colección de oligonucleótidos con posiciones interrogantes entre las cepas salvaje o mutante que se encuentran listadas en la Tabla II. El chip de ADN incluye, además de los nucleótidos que incluyen las 362 posiciones interrogantes listadas en la tabla (181 correspondientes al virus salvaje y otras 181 correspondientes a los mutantes de resistencia), todas las posibles variantes nucleotídicas individuales de cada uno de los codones en los que se ha descrito alguna mutación de resistencia (10 oligonucleótidos por posición, según lo indicado en la Tabla I). Las secuencias flanqueantes necesarias han sido diseñadas tanto para los oligonucleótidos salvaje como para los oligonucleótidos mutantes en función de su total homología de secuencia con la cepa HXB2 de VIH-1, subtipo B (Ratner *et al.*, 1985, número de acceso en la base de datos GenBank K03455).

Tales secuencias flanqueantes varían en función del tipo de chip de ADN diseñado. En el caso de utilizar OM1, las secuencias flanqueantes comprenden entre 5 y 50 nucleótidos a cada uno de los lados. Si se emplea la estrategia de hibridación por apilamiento de bases, la posición discriminadora está en el extremo 5' de un oligonucleótido de entre 5 y 100 nucleótidos de longitud. Para detectar genomas mutantes con inserciones de uno o dos aminoácidos, se diseñan oligonucleótidos mutantes que contienen los 3 ó 6 nucleótidos insertados, tal como se indica en la Tabla II.

Por otra parte, el contexto de secuencia en el que se encuentra la posición interrogante es variable en función de la diversidad genética del VIH. Como consecuencia de tal heterogeneidad, el VIH se clasifica en dos especies diferentes (VIH-1 y VIH-2) en grupos (M, O y N para VIH-1) y en subtipos (A-J para VIH-1 grupo M), cada uno con diferente distribución geográfica. Por ello, se diseñaron oligonucleótidos cuyas secuencias flanqueantes corresponden a subtipos distintos del B de VIH-1 grupo M (McCutchan *et al.*, 1996; Gao *et al.*, 1998; Paraskevis *et al.*, 1999), del VIH-1 grupo O (Janssens *et al.*, 1999; Mas *et al.*, 1999), al VIH-1 grupo N (Simon *et al.*, 1998) y al VIH-2 (Clavel *et al.*, 1986; Chen *et al.*, 1997). Para el análisis de virus recombinantes inter-subtipo o inter-grupo, se emplean oligonucleótidos diseñados en función de la secuencia que el recombinante posee en las regiones de PR y RT del gen *pol* (Robertson *et al.*, 1995; Takehisa *et al.*, 1999).

Otro de los motivos por los que se requiere una continua actualización en las secuencias de los oligonucleótidos es la constante descripción de nuevas mutaciones de resistencia, como respuesta al tratamiento con nuevos fármacos, o a nuevas terapias combinadas (Menéndez-Arias *et al.*, 1998; Winters *et al.*, 1998; Schinazi *et al.*, 1999; Briones *et al.*, 2000). La descripción de nuevas mutaciones también se debe en ocasiones al análisis de bases de datos que correlacionan patrones de resistencia genotípica y fenotípica (Hertogs *et al.*, 2000).

Por todo ello, debido a la constante detección de variantes genéticas de HIV ue difieren en mayor o menor grado con respecto a las secuencias conocidas, los catálogos de oligonucleótidos empleados son continuamente actualizables en función de las descripciones bibliográficas y las secuencias publicadas y/o depositadas en las bases de datos (Korber B *et al.*, 1998; actualizaciones periódicas en <http://hiv-web.lanl.gov>).

ES 2 273 835 T3

TABLA II

Listado de mutaciones de resistencia del Virus de la Inmunodeficiencia Humana (VIH) a inhibidores de la transcriptasa inversa (RT) y la proteasa (PR)

5

A) MUTACIONES ASOCIADAS A INHIBIDORES DE RT ANÁLOGOS DE NUCLEÓSIDO (NRTI)

NI ID:	Mutación		Aparece como mutación		Observaciones
	<i>wt</i>	Mut	Individual	Combinada	
10 1: M41L	ATG	TTG	+	-	M. 1 (AZT)
2: M41L	ATG	CTG	+	·	M. 1 (AZT); poco frec.
15 3: E44D	GAA	GAT	-	+	Poco frec.
4: E44A	GAA	GCA	-	+	Poco frec.
5: I50T	ATC	ACC	-	+	Poco frec.
20 6: A62V	GCT	GTT	-	+	M. 2 a Q151M (MDR)
7: K65R	AAA	AGA	·	+	M. 1 (ADV)
8: D67G	GAC	GGC	-	+	Poco frec.
25 9: D67N	GAC	AAC	-	·	
10: T69A	ACT	GCT	-	+	
30 11: T69D	ACT	GAT	+	·	M. 1 (ddC)
12: T69S	ACT	AGT	-	·	M. previa a la inserción de 2 aa
13-17:	-	i.: AGTAGT	-	+	Ins. de 2 aa (MDR)
35 69-SS-70		i.: AGTTCT			
		i.: AGCAGT			
		i.: AGCTCT			
40 18: 69-SG-70	-	i.: TCTAGT	-	+	Ins. de 2 aa (MDR)
19-20:	-	i.: AGTGGT	-	+	Ins. de 2 aa (MDR)
45 69-SA-70		i.: AGTGCT			
		i.: AGCGCT			
21: 69-ST-70	-	i.: TCTACC	-	·	Ins. de 2 aa: poco frec. (MDR)
50 22: 69-SV-70	-	i.: AGCGTG	-	·	Ins. de 2 aa: poco frec. (MDR)
23: 69-AG-70	-	i.: GCTGGT	-	·	Ins. de 2 aa: poco frec. (MDR)
24: 69-EA-70	-	i.: GAAGCA	-	+	Ins. de 2 aa: poco frec. (MDR)
55 25: 69-EE-70	-	i.: GAAGAA	-	·	Ins. de 2 aa: poco frec. (MDR)
26: 69-MT-70	-	i.: ATGACC	-	+	Ins. de 2 aa: poco frec. (MDR)
27: 69-IS-70	-	i.: ACCTCT	-	+	Ins. de 2 aa: poco frec. (MDR)

60

65

ES 2 273 835 T3

	28: 69-VG-70	-	i.: GTGGGT	-	-	Ins. de 2 aa: poco frec. (MDR)
	29: 69-D-70	-	i.: GAT	-	-	Ins. de 1 aa: poco frec. (MDR)
5	30: K70E	AAA	GAA	+	+	M. 1 (AZT); alto polimorfismo
	31: K70N	AAA	AAT	-	+	M. 1 (AZT); alto polimorfismo
10	32: K70N	AAA	AAC	-	+	M. 1 (AZT); alto polimorfismo
	33: K70R	AAA	AGA	+	+	M. 1 (AZT); alto polimorfismo
	34: L74V	TTA	GTA	+	-	M. 1 (ddl)
15	35: V75I	GTA	ATA	-	+	M. 2 a Q151M (MDR)
	36: V75T	GTA	ACA	+	+	M. 1 (d4T)
	37: F77L	TTC	CTC	-	+	M. 2 a Q151M (MDR)
20	38: E89G	GAA	GGA	-	-	Poco frec.
	39: V90I	GTT	ATT	-	+	Poco frec.
	40: A114S	GCT	AGT	-	+	Poco frec.
25	41: Y115F	TAT	TTT	+	+	Poco frec.
	42: F116Y	TTC	TAC	-	+	M. 2 a Q151M (MDR)
30	43: V118I	GTA	ATA	-	+	Poco frec.
	44: P119S	CCC	TCC	-	-	Poco frec.
	45: Q151M	CAG	ATG	-	+	MDR
35	46: P157S	CCG	TCG	-	-	
	47: R172K	AGA	AAA	-	-	Poco frec.
	48: I178M	ATA	ATG	-	+	Poco frec.
40	49: V179D	GTT	GAT	-	+	Poco frec.
	50: M184I	ATG	ATA	+	+	
	51: M184T	ATG	ACG	-	+	
45	52: M184V	ATG	GTG	+	+	M. 1 (3TC, ABC)
	53: L210W	TTG	TGG	-	+	
50	54: R211K	CGA	AAA	-	+	
	55: T215C	ACC	TGC	-	-	M. 1 (AZT); poco frec.
	56: T215F	ACC	TTC	+	-	M. 1 (AZT)
55	57: T215S	ACT	TCT	-	+	M. 1 (AZT); poco frec.
	58: T215Y	ACC	TAC	+	+	M. 1 (AZT)
	59: K219E	AAA	GAA	-	+	Poco frec.
60	60: K219Q	AAA	CAA	-	-	Poco frec.
	61: G333E	GGG	GAG	-	+	Poco frec.

65

ES 2 273 835 T3

B) MUTACIONES ASOCIADAS CON ANALOGOS NO-NUCLEOSIDOS DE INHIBIDORES DE RT (NNRTI)

5	NI ID:	Mutación		Aparece como mutación		Observaciones
		<i>wt</i>	Mut	Individual	Combinada	
10	1: E6K	GAG	AAG	-	+	Poco frec.
	2: L74I	TTA	ATA	-	+	Poco frec.
	3: L74V	TTA	GTA	-	+	Poco frec.
15	4: V75I	GTA	ATA	-	+	Poco frec.
	5: V75L	GTA	TTA	-	+	Poco frec.
	6: V90I	GTA	ATA	-	+	Poco frec.
20	7: A98G	GCA	GGA	+	-	
	8: L100I	TTA	ATA	+	-	
25	9: L100I	CTA	ATA	-	-	Menos frecuente que la anterior
	10: L100I	TTG	ATA	-	+	Menos frecuente que la anterior
	11: K101A	AAA	GCA	-	+	Poco frec.
30	12: K101E	AAA	GAA	+	+	Poco frec.
	13: K101I	AAA	ATA	-	+	Poco frec.
	14: K101Q	AAA	CAA	-	+	Poco frec.
35	15: K103N	AAA	AAC	+	+	M. 1 (NVP,EFV,DLV)
	16: K103N	AAG	AAT	+	+	M. 1 (NVP,EFV,DLV): poco frec.
	17: K103Q	AAA	CAA	-	+	Poco frec.
40	18: K103R	AAA	AGA	-	+	Poco frec.
	19: K103T	AAA	ACA	-	+	
45	20: V106A	GTA	GCA	+	+	
	21: V106I	GTA	ATA	-	+	Poco frec.
	22: V106L	GTA	TTA	-	+	Poco frec.
50	23: V108I	GTA	ATA	+	+	
	24: V108I	GTT	ATT	+	+	Menos frecuente que la anterior
	25: E138G	GAG	GGG	-	+	Poco frec.
55	26: E138K	GAG	AAG	-	+	Poco frec.
	27: E138R	GAG	AGG	-	+	Poco frec.
	28: T139I	ACA	ATA	-	+	Poco frec.
60	29: G141E	GGG	GAG	-	+	Poco frec.
	30: V179D	GTT	GAT	+	+	
	31: V179E	GTT	GAG	-	+	Poco frec.
65	32: Y181C	TAT	TGT	+	+	M. 1 (NVP,DLV)

ES 2 273 835 T3

	33: Y181H	TAT	CAT	-	+	Poco frec.
	34: Y181I	TAT	ATT	+	+	M. 1 (NVP,DLV); poco frec.
5	35: Y181L	TAT	CTT	-	-	Poco frec.
	36: Y188C	TAT	TGT	+	+	
	37: Y188H	TAT	CAT	-	+	Poco frec.
10	38: Y188L	TAT	TTA	+	+	
	39: Y188I	TAT	CTT	-	+	Poco frec.
	40: V189I	GTA	ATA	-	+	Poco frec.
15	41: G190A	GGA	GCA	+	+	
	42: G190E	GGA	GAA	-	-	Poco frec.
20	43: G190Q	GGA	CAA	-	+	Poco frec.
	44: G190S	GGA	AGC	+	+	
	45: G190T	GGA	ACA	-	+	Poco frec.
25	46: P225H	CCT	CAT	+	+	
	47: F227L	TTC	TTA	-	+	Poco frec.
	48: F227L	TTC	TTG	-	-	Poco frec.
30	49: F227L	TTC	CTC	-	+	Poco frec.
	50: M230L	ATG	TTG	-	-	Poco frec.
35	51: E233V	GAA	GTA	-	-	Poco frec.
	52: L234I	CTC	ATC	-	+	Poco frec.
	53: P236L	CCT	CTT	-	+	
40	54: K238T	AAA	ACA	-	-	Poco frec.

C) MUTACIONES ASOCIADAS CON INHIBIDORES DE PROTEASA (PI)

	NI ID:	Mutación		Aparece como mutación		Observaciones
		WT	Mut	Individual	Combinada	
50	1: R8Q	CGA	CAA	+	-	
	2: R8K	CGA	AAA	-	+	Poco frec.
55	3: L10F	CTC	CGC	-	+	
	4: L10F	CTC	TTC	-	-	Menos frecuente que la anterior
	5: L10I	CTC	ATC	-	+	
60	6: L10R	CTC	CGC	-	+	
	7: L10V	CTA	GTA	-	-	Poco frec.
	8: L10Y	AAA	TAT	-	+	Poco frec.
65	9: L10Y	AAA	TAC	-	+	Poco frec.

ES 2 273 835 T3

	10: I11V	ATA	GTA	-	+	Poco frec.
	11: I13V	ATA	GTA	-	+	
5	12: K20M	AAG	ATG	-	+	
	13: K20R	AAG	AGG	-	+	
	14: L23I	CTA	ATA	-	+	Poco frec.
10	15: L23V	TTA	GTA	-	+	
	16: L24I	TTA	ATA	-	+	
15	17: L24V	TTA	GTA	-	+	Poco frec.
	18: D30N	GAT	AAT	+	·	M. 1 (NFV)
	19: V32I	GTA	ATA	-	+	
20	20: L33F	TTA	TTT	-	+	
	21: F34K	GAA	AAA	-	+	Poco frec.
	22: E34V	GAA	GTA	-	+	Poco frec.
25	23: F35D	GAA	GAT	-	+	
	24: M36I	ATG	ATA	-	+	M. 2 muy frecuente
	25: K45E	AAA	GAA	-	+	Poco frec.
30	26: K45I	AAA	ATA	-	+	Poco frec.
	27: M46F	ATG	TTC	-	+	Poco frec.
35	28: M46I	ATG	ATA	+	+	M. 1 (IDV)
	29: M46L	ATG	TTG	+	-	M. 1 (IDV)
	30: M46V	ATG	GTG	-	+	Poco frec.
40	31: I47A	ATA	GCA	-	+	Poco frec.
	32: I47V	ATA	GTA	-	+	
	33: G48V	GGG	GTG	+	+	M. 1 (SQV)
45	34: I50L	ATC	TTA	-	+	Poco frec.
	35: I50I	ATC	CTC	-	-	Poco frec.
	36: I50V	ATT	GTT	-	+	M. 1 (APV)
50	37: I54I	ATC	TTA	-	+	Poco frec.
	38: I54L	ATC	CTC	-	+	Poco frec.
55	39: I54M	ATC	ATG	-	+	
	40: I54V	ATC	GTC	-	-	M. 2 muy frecuente
	41: D60E	GAT	GAA	-	+	Poco frec.
60	42: L63P	CTC	CCC	-	-	Alto polimorfismo natural
	43: L63Q	CTG	CAG	-	+	Alto polimorfismo natural
	44: L63V	TTA	GTA	-	·	Alto polimorfismo natural
65	45: A71T	GCT	ACT	-	+	

ES 2 273 835 T3

	46: A71V	GCT	GTT	-	+	M. 2 muy frecuente
	47: G73S	GGT	GCT	-	+	
5	48: G73S	GGT	AGT	-	+	Menos frecuente que la anterior
	49: V75I	GTC	ATC	-	+	Poco frec.
	50: I.76M	TTG	ATG	-	·	Poco frec.
10	51: V77I	GTC	ATC	-	-	
	52: P81T	CCT	ACT	-	-	Poco frec.
15	53: V82A	GTC	GCC	·	+	M. 1 (IDV,RTV)
	54: V82F	GTC	TTC	+	+	M. 1 (IDV,RTV)
	55: V82I	GTC	ATC	+	+	M. 1 (IDV)
20	56: V82S	GTC	TCC	+	+	M. 1 (RTV)
	57: V82T	GTC	ACC	+	·	M. 1 (IDV,RTV)
	58: I84A	ATA	GCA	-	·	Poco frec.
25	59: I84V	ATA	GTA	-	+	M. 2 muy frecuente
	60: N88D	AAT	GAT	-	·	
	61: N88S	AAT	AGT	-	+	
30	62: I.89M	TTG	ATG	-	+	
	63: I.90I	TTA	ATA	-	+	Poco frec.
35	64: I.90M	TTG	ATG	·	+	M. 1 (SQV,NFV)
	65: T91S	ACT	TCT	-	+	Poco frec.
	66: L97V	TTA	GTA	-	+	Poco frec.

40 *Leyenda de la Tabla II*

1. Las mutaciones de resistencia se han dividido con respecto a las tres familias de fármacos antirretrovirales: (A) inhibidores de RT análogos de nucleósido (NRTI), (B) inhibidores de RT no análogos de nucleósido (NNRTI) y (C) inhibidores de la proteasa (PI). Estas posiciones son las que se han descrito como responsables de la resistencia a distintos fármacos. En el microchip de ADN específico de VIH se incluyen todas ellas, junto con todas las posibles variantes nucleotídicas de cada uno de los codones contemplados.

2. La primera columna muestra un número de identificación (Nº ID) y el cambio de aminoácido asociado a la resistencia. Se indica el aminoácido de la cepa salvaje (wt) la posición que éste ocupa en el gen, y el aminoácido de la cepa mutante. Así, por ejemplo, el cambio M41L en la parte (A) de la tabla corresponde a un cambio de metionina por leucina en la posición 41 del gen de la RT. Las mutaciones con número de identificación 13 a 29 de la parte (A) corresponden a inserciones de los aminoácidos indicados entre los codones 69 y 70 de la RT.

3. Las columnas segunda y tercera muestran la secuencia nucleotídica de las cepas salvaje (wt) y mutante (mut). En los casos de inserción de aminoácidos, la columna correspondiente al genoma mutante muestra los 3 ó 6 nucleótidos insertados. Como genoma salvaje de referencia se utiliza la cepa HXB2 de VIH-1 subtipo B.

4. La sexta columna recoge una serie de observaciones y datos adicionales relacionadas con cada mutación. M1 y M2 indican respectivamente mutaciones primarias o secundarias para el inhibidor, lo que implica que se desarrollen más pronto o más tarde como respuesta al tratamiento con dicho fármaco. Se señala también si la mutación aparece con poca frecuencia entre los pacientes tratados (poco frec.), si consiste en una inserción de aminoácidos (Ins.) o si esa posición resulta especialmente variable en VIH-1, presentando por tanto un alto polimorfismo natural. En los casos en que los fármacos responsables de cada mutación se indican entre paréntesis, se han empleado las siguientes abreviaturas: AZT, zidovudina; ddI, didanosina; ddC, zalcitabina; 3TC, lamivudina; d4T, estavudina; ABC, abacavir; ADV, adefovir; NVP, nevirapina; EFV, efavirenz; DLV, delavirdina; SQV, saquinavir; RTV, ritonavir; IDV, indinavir; NFV, nelfinavir; APV, amprenavir. En la parte (A) se indican como MDR las mutaciones asociadas con multiresistencia a diversos fármacos de una misma familia.

ES 2 273 835 T3

5. Los datos recogidos en la Tabla II proceden de De Antoni *et al.*, 1997; Winters *et al.*, 1998; Menéndez-Arias *et al.*, 1998; Schinazi *et al.*, 1999; De Jong *et al.*, 1999; Korber *et al.*, 1998; Briones *et al.*, 2000; y Hertogs *et al.*, 2000.

Se construyó un microchip de ADN con oligonucleótidos memoria de tipo 1 con una longitud de 15 nt, complementarios a la región del genoma viral que incluye los codones de la Tabla II. Todos ellos pertenecen a los genes de la proteasa (PR) o de la transcriptasa inversa (RT) del HIV-1 HXB2 subtipo B. En este genoma tipo, los fragmentos de la proteasa (PR) y de la transcriptasa inversa (RT) del gen *pol* corresponden, respectivamente, a las posiciones 2253-2549, y 2550-4229.

Se llevaron controles negativos adicionales como los oligonucleótidos universales de pUC18 (Sambrook *et al.* 1989) y otros que no se esperan encontrar en la población viral como controles negativos. Se procedió como sigue:

1. Extracción de ARN viral y síntesis de cADN

Las muestras de sangre se extrajeron de los pacientes VIH-positivos en tubos de 10 ml con EDTA. La separación del plasma a partir de la sangre total se realiza mediante centrifugación a 5.000 g durante 20 minutos. Con ello se obtiene una fase superior que contiene el plasma, una fase inferior (hematocrito), y una interfase que contiene los linfocitos de sangre periférica (PBMCs).

A partir de 0,5-1 ml del plasma obtenido, el ARN de VIH se obtiene mediante lisis viral con isotiocionato de guanidina, seguida de adsorción a partículas de sílice, lavado y resuspensión (Boom *et al.*, 1990). Un método alternativo para la extracción de ARN viral consiste en la ultracentrifugación del plasma a 23.000 g durante 1 hora para obtener un sedimento de partículas virales, seguida de lisis de los virus y resuspensión del ARN.

2. Extracción de ADN proviral

En los casos en que se analiza en ADN proviral integrado en el genoma de los linfocitos de sangre periférica (PBMCs), tras la separación de las células tal como se indica en el punto 1 se procede a la extracción del ADN. El sistema empleado consiste en la sedimentación de los PBMCs por centrifugación a 10.000 g durante 5 min, resuspensión y lisis celular mediante proteinasa K/Tween 20 (Innis *et al.*, 1990).

3. Síntesis de cADN

La síntesis de cADN a partir del ARN viral obtenido según lo indicado en el punto 1 se realiza como se describe a continuación: 10 ml de ARN se incuban con 300 ng de un iniciador específico de la región que se va a retrotranscribir (iniciador B para la RT o iniciador F para la PR, ver puntos 4 y 5). Tras una incubación a 70°C durante 10 minutos, se añade 8 ml de una mezcla que contiene: 10 unidades de AMV-RT (Promega, Madison, MI), 1 ml de mezcla de dNTPs 10 mM, 25 unidades de RNAsina, 4 ml de Buffer de AMV-RT 5x y 1,37 ml de agua. Se realiza una incubación a 42°C durante 1 hora seguida de otra a 70°C durante 5 minutos. El cADN puede ser utilizado directamente para realizar amplificación por PCR, o mantenido a 70°C hasta el momento de su uso.

4. Amplificación por PCR de un fragmento de la RT

Mediante PCR anidada se amplifica un fragmento de la RT de 647 bp (codones 19 a 234). La primera ronda de PCR se realiza utilizando el iniciador directo A y el reverso B:

A: 5'-GGTTGCACTTTAAATTTTCCCATTAGTCCTATT-3'

B: 5'-TACTAACTTCTGTATGTCATTGACAGTCCAGCT-3'

La segunda ronda de PCR se lleva a cabo empleando los iniciadores C y D:

C: 5'-GTAAACAATGGCCATTGACAG-3'

D: 5'-AGTTCATAACCCATCCAAAGG-3'

Las condiciones empleadas son:

a) para la primera PCR: una desnaturalización inicial de 5 minutos a 94°C seguida de 40 ciclos de (94°C-30 s/55°C-30 s/72°C-1 minuto) y una elongación final de 5 minutos a 72°C; y

b) para la segunda PCR: desnaturalización de 1 minuto a 94°C, 35 ciclos de (94°C-30 s/57°C-30 s/72°C-30 s) y elongación final de 5 minutos a 72°C.

5. Amplificación por PCR de la PR

Mediante PCR anidada se amplificaron 401 pares de bases conteniendo la totalidad de la PR. La primera ronda de PCR se realizó utilizando el iniciador directo E y el reverso F:

ES 2 273 835 T3

E: 5'-GCCAACAGCCCCACCAGAAGAGAGC-3'

F: 5'-GGCCATTGTTTAACTTTTGGGCCATCC-3'

5 La segunda ronda de PCR se llevó a cabo empleando los iniciadores G y H:

G: 5'-CAACTCCCTCTCAGAAGCAGGAGCCG-3'

H: 5'-CCATTCTGGCTTTAATTTTACTGGTA-3'

10

Las condiciones empleadas son:

a) para la primera PCR: una desnaturalización inicial de 5 minutos a 94°C seguida de 35 ciclos de (94°C-30 s/56°C-30 s/72°C-1 minuto) y una elongación final de 5 minutos a 72°C; y

15

b) para la segunda PCR se emplean las mismas condiciones, salvo que la fase de elongación en los ciclos es de 54°C durante 30 s.

20 6. Amplificación por PCR de la PR y RT completas

Es posible amplificar en un solo proceso de PCR anidada la totalidad de las regiones PR y RT del gen *pol*. No obstante, dado que el fragmento generado es muy largo (más de 2.000 nucleótidos), la eficiencia de la amplificación disminuye notablemente respecto a las amplificaciones de PR y RT por separado. Dado que en el método objeto de la presente invención se va en busca de genomas memoria minoritarios dentro de la cuasiespecie, resulta adecuado optimizar la eficiencia de la amplificación. Por ello, es preferible realizar amplificaciones de las regiones de interés tal como se indica en los puntos 4 y 5.

25

7. Purificación de los fragmentos amplificados por PCR y marcaje fluorescente

Los fragmentos del genoma viral amplificados por PCR se aíslan por fraccionamiento en geles de agarosa y se purifican a través de columnas QIAquick (QIAquick PCR purification kit, Qiagen #28106). Aproximadamente 1 mg de cada uno de los fragmentos purificados se mezcla en un tubo y se marca fluorescentemente en la siguiente mezcla de reacción: 71 μ l de H₂O, 10 μ l de Buffer PCR 10X (500 mM KCl, 100 mM Tris-Cl pH 8,3, 15 mM MgCl₂, 0,1% gelatina), 10 μ l de dNTPs (2 mM cada uno), 5 μ l Cy5-dCTP (1 mM), 2,0 μ l (100 pmol/ μ l) de los oligonucleótidos adecuados para RT y para PR, 1 μ g (1-2 ml) de la mezcla de ADN amplificado por PCR, 1,0 μ l de *Taq* ADN polimerasa (5 unidades/ μ l). Se generan cADNs fluorescentes de cadena sencilla por amplificación lineal del molde según las siguientes condiciones: desnaturalización a 95°C durante 2 minutos, amplificación durante 30 ciclos de (94°C-30 s/55°C-30 s/72°C-30 s), extensión final a 72°C durante 3 minutos. Los productos fluorescentes se purifican a través de una columna de QIAquick (Qiagen) y se resuspenden en 50 μ l de 1X TE (10 mM Tris-Cl y 1 mM EDTA) pH 8,0, hasta una concentración final aproximada de 20 ng/ μ l. Por cada 13 μ l se añaden 5,0 μ l de SSC 20 X [3 M NaCl, 0,3 M citrato sódico (pH 7,0)] y 2,0 μ l 2% SDS, se calienta a 65°C durante 30 s para disolver la muestra fluorescente, y tras una centrifugación de 2 minutos a 5.000 g para eliminar sedimentos, se transfiere el sobrenadante a un tubo nuevo. La concentración final de la mezcla fluorescente es aproximadamente de 0,15 μ g/ μ l, en 20 μ l de 5XSSC y 0,2% de SDS.

35

40

Se procede de igual forma para el marcaje con un fluorocromo distinto (Cy3) de un fragmento de ADN amplificado con los mismos oligonucleótidos para la RT y para la PR del virus VIH tipo.

45

8. Diseño y construcción del chip de oligonucleótidos memoria del HIV

La secuencia mayoritaria en la RT y PR de la cuasiespecie viral extraída del paciente se determina mediante la secuenciación de las dos cadenas de cada fragmento en la cuasiespecie viral total (secuencia consenso), usando los mismos oligonucleótidos que para la amplificación interna por PCR anidada.

50

Sobre un portaobjetos de vidrio de microscopía preparado con un aldehído activo (CEL Associates) se imprimen 600 oligonucleótidos OM1 de 15 nt de longitud, donde la posición central (la octava) es la posición interrogante, y por tanto identifica la mutación en la muestra marcada fluorescentemente. En cada punto del chip se depositaron 2 pmol (picomoles) de oligonucleótido en solución de impresión y se eliminaron los oligonucleótidos no incorporados. La disposición de los oligonucleótidos en el chip es la siguiente:

55

a) un juego de OM1 de secuencia igual a la del genoma VIH-1, subtipo B (Ratner *et al.*, 1985, número de acceso GenBank K03455) por duplicado, cada oligonucleótido con 7 nucleótidos flanqueantes a cada lado de cada una de las mutaciones puntuales enumeradas en la Tabla II;

60

b) un juego de OM1 de secuencia igual a la secuencia mayoritaria encontrada en una muestra de secuencia conocida procedente de un paciente, también por duplicado. De nuevo la secuencia de estos OM1 comprende 7 nucleótidos flanqueantes a cada lado de cada una de las mutaciones puntuales enumeradas en la Tabla II. Este juego de oligonucleótidos constituye un control positivo, de tal forma que la intensidad de la señal de hibridación identifica o es característico del nivel del genoma mayoritario;

65

ES 2 273 835 T3

c) un juego de OM1 (por duplicado) de secuencia igual a la del juego del apartado b) excepto en la posición interrogante donde ahora está la base que identifica cada una de las mutaciones enumeradas en la Tabla II;

5 d) una serie por cuadruplicado de mezclas a distintas proporciones de los oligonucleótidos salvaje (wt) y mutante (mut) para el aminoácido 41 de la RT (M41L). Las proporciones del mutante respecto al salvaje en la mezcla son 1, 5×10^{-1} , 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 0. La representación gráfica de la media de la intensidad de la señal de hibridación en cada uno de estos puntos nos permite construir una curva patrón para cuantificar las señales de hibridación correspondientes a mutaciones de genomas memoria;

10 e) una serie de 50 puntos conteniendo la misma cantidad de oligonucleótido wt para el codón 41 de la RT (ATG);

f) una serie de 50 puntos conteniendo la misma cantidad de oligonucleótido mut para el codón 41 de la RT (ATG→TTG);

15 g) una serie de 50 puntos conteniendo la misma cantidad de oligonucleótido mut para el codón 41 de la RT (ATG→CTG);

20 h) una serie de 50 puntos conteniendo la misma cantidad de oligonucleótido mut para el codón 41 de la RT (ATG→GTG). Las series (e-h) son controles que nos dan información sobre la intensidad de la señal para la copia salvaje (e), y para las posibles mutaciones memoria (f-g). A su vez, una o las tres series de oligonucleótidos mutantes sirven de control negativo puesto que se espera que su secuencia complementaria no esté en la mezcla fluorescente;

25 i) dos series de 10 puntos cada una conteniendo los oligonucleótidos universales del plásmido pUC18, que no tienen secuencias complementarias a ningún oligonucleótido del VIH. Estos puntos sirven como control negativo de hibridación 0%.

9. Hibridación de las muestras fluorescentes con el microchip

30 Aproximadamente 1/5 del volumen de las mezclas fluorescentes se desnaturaliza hirviendo durante 2 minutos y se pone en contacto con el chip para permitir la hibridación durante 4 horas a 42°C. La eliminación de la muestra no hibridada se hace mediante dos lavados con SSC 2X a temperatura ambiente (22-25°C) durante 5 minutos cada uno, y un tercer lavado en las mismas condiciones.

10. Escaneado del microchips

35 La disponibilidad de un escáner que permite la lectura del microchip húmedo (sin necesidad de secado previo del mismo) permite comprobar el resultado de cada lavado y modificar las condiciones para aumentar la relación señal/ruido de la hibridación. Una vez terminados los lavados, el microchip se deja secar al aire. El escaneado se realiza con los mayores valores del PMT (tubo fotomultiplicador) y del láser que conservan la linealidad y minimizan el ruido de fondo.

11. Análisis de las imágenes escaneadas

45 La intensidad de la fluorescencia se cuantificó con el software ImaGene. Existe un software variado disponible gratuitamente a través de internet. No obstante, como se indicó previamente, otro aspecto de la invención contempla el desarrollo de nuevos programas informáticos que mejorarán la precisión de los análisis.

50 Como ejemplo, se describe brevemente el resultado obtenido mediante utilización de la invención sobre una muestra clínica de un paciente VIH-positivo (véanse las abreviaturas de los fármacos en la leyenda de la Tabla II):

55 i) el paciente había estado recibiendo tratamiento con AZT+3TC+IDV durante 8 meses, y se determinó un fracaso de dicha pauta terapéutica en base a la elevación de la carga viral hasta un valor de 28.540 copias de ADN de VIH por ml de plasma sanguíneo y deterioro inmunitario medido por recuento de linfocitos CD4+. La aplicación del método de la invención en ese momento mostró las siguientes mutaciones asociadas con resistencia en el genoma mayoritario: M46L en la PR y M41L, M184V, L210W, T215Y en la RT;

60 ii) en este momento se realizó un cambio de pauta terapéutica con la siguiente combinación: ddI+d4T+IDV, que se mantuvo durante 5 meses hasta el segundo análisis del genotipo de la cuasiespecie viral empleando el método de la invención. Se detectaron las siguientes mutaciones asociadas con resistencia en el genoma mayoritario: M46L, L63P, A71V, I84V en la PR. Asimismo, se detectaron las siguientes mutaciones asociadas con resistencia en el genoma minoritario (en una proporción entre 10^{-3} y 5×10^{-2}): M46L y L63P en la PR; M41L y T215Y en la RT; y

65 iii) la interpretación de estos resultados indica que tras el tratamiento con ddI+d4T+IDV, ambas mutaciones presentes en la RT en genomas memoria minoritarios (pero no en el genoma mayoritario), han quedado como mutaciones de memoria de genomas que en el análisis anterior aparecían como mayoritarios. En concreto, estas dos mutaciones que condicionan resistencia a AZT, fueron seleccionadas como mayoritarias durante el tratamiento anterior (AZT+3TC+IDV) como respuesta a uno de los fármacos de la combinación terapéutica. Tras 5 meses con un tratamiento que no incluye AZT, las mutaciones de resistencia a ese fármaco se mantienen en genomas memoria. La

ES 2 273 835 T3

utilidad de la invención radica en que el desconocimiento de la existencia de estas mutaciones en los genomas memoria podría llevar al médico a prescribir AZT en una nueva pauta terapéutica puesto que no encuentra mutaciones a ese fármaco en el genoma mayoritario. Sin embargo, la presencia de tales mutaciones en genomas memoria de la cuasiespecie (como resultado de una terapia anterior que el médico puede no conocer) indica que podrían ser rápidamente seleccionadas ante una nueva exposición a AZT, lo que llevaría a un rápido fracaso de la terapia en dicho paciente.

Ejemplo 5

Método de detección de genomas memoria en cuasiespecies virales del virus de la hepatitis C (VHC)

Se puede interrogar cualquiera de las posibles mutaciones en las regiones del VHC comprendidas entre los nucleótidos 6967-7086 (región ISDR, por Interferon Sensitive Determining Region), 2319-2351 (tomando como nucleótido 1 el primer nucleótido del genoma viral) genotipo 1-b y las zonas (+/-20 nt) alrededor de las posiciones 195 y 330, donde las dos primeras corresponden a las regiones relacionadas con respuesta a interferón (IFN) y las dos últimas con actividad ribozimática. La ISDR está directamente implicada en la respuesta a interferón (N.Engl.J.Med.334:77-81), mientras que la región 1978-2010 se asocia con un posible mecanismo de interferencia en la acción del IFN a través de la PKR (Science 285:107-109). En la ISDR, la resistencia al IFN no parece que este mediada por la selección de ninguna mutación particular, sino que los pacientes infectados por una secuencia parecida a una determinada variante responden con peor resultado al IFN que las secuencias que tienen varias mutaciones de esta variante. Con respecto a los ribozimas, no hay datos clínicos aún, aunque algunos ya están siendo estudiados en ensayos clínicos de fase uno. Aunque no existen todavía evidencias concluyentes relativos al efecto de mutaciones puntuales en estas regiones del VHC sobre la resistencia a determinados fármacos, la detección de genomas memoria minoritarios podría aportar datos de enorme interés sobre la persistencia del virus en un organismo infectado. Por ejemplo, como se muestra en la Figura 1, existen diferencias significativas entre la secuencia de los genotipos 1-a y 1-b para las regiones implicadas en respuesta a INF.

El procedimiento seguido es:

1) El aislamiento, purificación y análisis del ARN viral se lleva a cabo según Cabot *et al.*, 2000 y Gerotto y col., 1999, y los kits con los que se extrae el ARN del VHC son de ROCHE (referencias: 2065193 y 2065231000).

2) Amplificación por RT-PCR y/o PCR anidada de los fragmentos de ADN conteniendo las zonas objeto de estudio, con los juegos de oligonucleótidos citados en la Tabla II y usando como molde los ácidos nucleicos extraídos según la etapa 1. Los fragmentos amplificados contienen las regiones del genoma del VHC tipo 1-b (respecto del primer nucleótido del genoma viral) entre las posiciones 6967-7086 (región ISDR), 2319-2351, 175-215 (región ribozima 1) y 310-350 (región ribozima 2).

3) Clonaje de los fragmentos en un vector bacteriano y determinación de la secuencia del clon mayoritario.

4) Amplificación por PCR de los fragmentos clonados y marcaje de los fragmentos amplificados con 32P- γ -ATP y polinucleótido quinasa de T4.

5) Hibridación de los fragmentos marcados con los fragmentos amplificados directamente del ácido nucleico de la cuasiespecie, y consigo mismo sin marcar.

6) Fraccionamiento de los híbridos formados por electroforesis en un gel de poliacrilamida en condiciones nativas. El fragmento marcado consigo mismo sirve de control.

7) Identificación de la existencia de genomas minoritarios por el número de mutaciones en función del salto en la movilidad electroforética.

8) Extracción de los ADN hibridados del gel de poliacrilamida por elución. Amplificación por PCR de los fragmentos eluidos, secuenciación de los fragmentos amplificados y comparación de las secuencias de nucleótidos deducidas.

Ejemplo 6

Método de detección de genomas memoria del virus de la hepatitis B (VHB) portadores de mutaciones que confieren resistencia a fármacos en enfermos tratados

Siguiendo el procedimiento general descrito en el Ejemplo 2, con las modificaciones oportunas, se diseña un microchip de ADN que contiene todas las posiciones nucleotídicas codificantes del fragmento comprendido entre los aminoácidos 500 y 600 de la transcriptasa inversa (RT) de VHB. Esta región contiene el "motivo YMDD" implicado en la resistencia del VHB a lamivudina (3TC) (Allen *et al.*, 1998). Los resultados permiten detectar la presencia de genomas memoria minoritarios con el cambio de aminoácido M552V (cambio nucleotídico ATG→GTG) en pacientes que habían estado sometidos a tratamiento con 3TC.

REIVINDICACIONES

1. Un método para diseñar una terapia antiviral personalizada para un sujeto basada en la detección de genomas memoria minoritarios en una cuasiespecie viral responsable de la resistencia viral a un fármaco o combinación de fármacos, que comprende:

a) detectar genomas memoria minoritarios en una cuasiespecie viral de una muestra sospechosa de contener dichos genomas memoria minoritarios en donde dichos genomas memoria minoritarios están presentes en una proporción desde 10^{-5} (0,001%) a menos del 50% de dicha cuasiespecie viral, contienen al menos una mutación en comparación con el genoma mayoritario de dicha cuasiespecie viral, y proporcionan información sobre la historia evolutiva de dicha cuasiespecie viral en dicho sujeto,

b) detectar la existencia de mutaciones nucleotídicas asociadas con la resistencia a fármacos antivirales en dichos genomas memoria minoritarios, y

c) considerar los datos obtenidos en las etapas previas para diseñar una terapia antiviral personalizada, de forma que emplea uno o más fármacos antivirales contra los cuales dicha cuasiespecie viral no presenta resistencia basada en mutaciones en dichos genomas memoria minoritarios.

2. Método según la reivindicación 1, en donde la detección de dichos genomas memoria minoritarios en una cuasiespecie viral, comprende:

a) extraer el ácido nucleico de una muestra sospechosa de contener dichos genomas memoria minoritarios en una cuasiespecie viral;

b) amplificar al menos un fragmento de ácido nucleico de dicha cuasiespecie viral; y

c) detectar y analizar la existencia de genomas memoria minoritarios usando técnicas seleccionadas entre el uso de microchips de ADN, ensayo de la huella del heteroduplex y clonaje molecular.

3. Método según la reivindicación 2, que comprende:

a) extraer el ácido nucleico de una muestra sospechosa de contener dichos genomas memoria minoritarios en una cuasiespecie viral;

b) amplificar al menos un fragmento de ácido nucleico de dicha cuasiespecie viral;

c) marcar el fragmento o fragmentos amplificados con un compuesto marcador;

d) construir un microchip de ADN que comprende:

i) al menos un oligonucleótido que sirve como un control positivo,

ii) al menos un oligonucleótido que sirve como un control negativo,

iii) al menos un oligonucleótido memoria, y

iv) unos medios que permiten realizar una curva de calibrado;

e) poner en contacto dichos fragmentos amplificados en la etapa b) y marcados en la etapa c) con los oligonucleótidos presentes en el microchip de ADN preparado en la etapa d) bajo condiciones que permiten la hibridación, únicamente cuando todos los nucleótidos de un oligonucleótido presente en dicho microchip de ADN aparean con una secuencia de nucleótidos presente en dichos fragmentos amplificados y marcados;

f) identificar los oligonucleótidos presentes en dicho microchip de ADN que han hibridado con dichos fragmentos amplificados y marcados; descartando las hibridaciones negativas o ruido de fondo; y

g) seleccionar los oligonucleótidos presentes en dicho microchip de ADN que han hibridado con dichos fragmentos amplificados y marcados y que, por interpolación con la curva de calibrado muestran una proporción de dichos fragmentos en la cuasiespecie inferior al 50% característico de genomas memoria minoritarios.

4. Método según la reivindicación 3, en el que dicho genoma memoria minoritario se encuentra en una proporción comprendida entre el 0,1% y el 10% de la cuasiespecie.

5. Método según la reivindicación 3, en el que dicha muestra sospechosa de contener dichos genomas memoria minoritarios en una cuasiespecie viral es una muestra seleccionada entre una muestra clínica o una derivada de un cultivo viral.

ES 2 273 835 T3

6. Método según la reivindicación 3, en el que dicha cuasiespecie viral pertenece al virus de la inmunodeficiencia humana tipo 1 (VIH-1).
7. Método según la reivindicación 3, en el que dicha cuasiespecie viral pertenece al virus de la inmunodeficiencia humana tipo 2 (VIH-2).
8. Método según la reivindicación 3, en el que dicha cuasiespecie viral pertenece al virus de la hepatitis C (VHC).
9. Método según la reivindicación 3, en el que dicha cuasiespecie viral pertenece al virus de la hepatitis B (VHB).
10. Método según la reivindicación 3, en el que dicha cuasiespecie viral pertenece al virus de la fiebre aftosa (VFA).
11. Método según la reivindicación 3, que comprende la realización de una transcripción inversa del ARN viral previa a la etapa b) de amplificación.
12. Método según la reivindicación 3, en el que dicha amplificación se realiza por métodos enzimáticos.
13. Método según la reivindicación 12, en el que dichos métodos enzimáticos comprenden la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR) o la amplificación basada en transcripción (TAS).
14. Método según la reivindicación 3, en el que el fragmento a amplificar en la etapa b) corresponde a una parte o a la totalidad de al menos un gen imprescindible para la replicación o la persistencia del virus en el organismo infectado.
15. Método según la reivindicación 14, en el que dicho gen imprescindible se selecciona del grupo comprendido por: el fragmento de la proteasa (PR) del gen *pol* de VIH, el fragmento de la transcriptasa inversa (RT) del gen *pol* de VIH, el fragmento de la integrasa del gen *pol* de VIH, el gen *env* de VIH, el gen *gag* de VIH, el gen de la proteína no estructural NS5A de VHC, la región comprendida entre los nucleótidos 175-215 de VHC, la región comprendida entre los nucleótidos 310-350 de VHC y el fragmento de la transcriptasa inversa (RT) del gen *pol* de VHB.
16. Método según la reivindicación 3, en el que dicho compuesto marcador utilizado para marcar los fragmentos amplificados se selecciona de un compuesto radiactivo, un compuesto fluorescente o un compuesto detectable por una reacción colorimétrica.
17. Método según la reivindicación 3, en el que dicho microchip de ADN está constituido por puntos de oligonucleótidos sintetizados previamente.
18. Método según la reivindicación 3, en el que dicho microchip de ADN está compuesto por puntos de oligonucleótidos previamente sintetizados *in situ*.
19. Método según la reivindicación 3, en el que dicho control positivo comprende al menos un oligonucleótido con al menos una secuencia de oligonucleótidos que es 100% complementaria a una secuencia conocida del genoma mayoritario o del genoma salvaje del virus.
20. Método según la reivindicación 3, en el que dicho control negativo se selecciona del grupo formado por: i) un oligonucleótido con una secuencia que es complementaria a una zona de secuencia conocida del genoma mayoritario o del genoma salvaje del virus, excepto en al menos una posición; ii) un oligonucleótido con una secuencia que es complementaria a una zona de secuencia conocida del genoma mayoritario o salvaje del virus, excepto en la posición interrogante; y iii) un oligonucleótido con una secuencia que es complementaria a una zona de secuencia conocida del genoma mayoritario o salvaje del virus, excepto en la posición interrogante y en al menos un nucleótido flanqueante de dicha posición interrogante.
21. Método según la reivindicación 3, en el que dicho oligonucleótido memoria se selecciona del grupo formado por:
- un ácido nucleico con una longitud de 4 a 250 nt que es igual o complementario a una secuencia del genoma viral mayoritario o promedio excepto en las posiciones centrales 1-6 (OM1);
 - un ácido nucleico de 5-50 nt en longitud que está formado por apilamiento de dos oligonucleótidos tras hibridar con otro ácido nucleico complementario procedente del virus, estando constituido uno de los oligonucleótidos estancados por una mezcla de cuatro oligonucleótidos que difieren en la posición inmediatamente adyacente al oligonucleótido anterior y que llevan un colorante fluorescente diferente unido covalentemente al otro extremo (OM2);
 - un ácido nucleico de entre 5 y 250 nt comprendido de dos partes, una 5' complementaria al otro oligonucleótido ausente del genoma viral, y una parte 3' complementaria al genoma viral, siendo la última posición una posición interrogante (OM3);
 - un ácido nucleico de 5 a 250 nt de longitud complementario al genoma viral que lleva una sustancia fluorescente unida covalentemente al extremo 3' (OM4);

ES 2 273 835 T3

- un ácido nucleico de 5 a 250 nt de longitud complementario al genoma viral cuya última posición del extremo 3' es anterior a una posición interrogada del genoma viral (OM5);

5 - un ácido nucleico de 5 a 250 nt de longitud complementario a una secuencia de un genoma mayoritario de una cuasiespecie viral con inserciones de 1 a 10 nt respecto a la secuencia del genoma mayoritario (OM6);

- un ácido nucleico entre 5 y 250 nt complementarios a una secuencia del genoma mayoritario de la cuasiespecie viral con deleciones 1 a 10 nt respecto a dicha secuencia del genoma mayoritario (OM7);

10 - un ácido nucleico de 5 a 250 nt complementarios a una secuencia mutante previamente descrita en las bases de datos; y

- sus mezclas.

15 22. Método según la reivindicación 3, en el que dicha curva de calibrado se elabora usando unas series de mezclas de oligonucleótidos en proporciones variables y conocidas, siendo uno de estos 100% complementario a una zona de secuencia conocida del genoma mayoritario o del genoma salvaje, y el otro un oligonucleótido que difiere del anterior en al menos una posición.

20 23. Método según la reivindicación 3, en el que dicha curva de calibrado está formada por series de mezclas de oligonucleótidos en proporciones variables y conocidas, siendo uno de ellos 100% complementario a una zona de secuencia conocida del genoma mayoritario o del genoma salvaje, y el otro un oligonucleótido que difiere del anterior en la posición interrogante.

25 24. Método según la reivindicación 3, en el que la identificación de los oligonucleótidos presentes en el microchip de ADN que han hibridado con los fragmentos amplificados y marcados se realiza mediante escaneado de dicho microchip con un escáner provisto de un microscopio confocal y al menos dos láseres que emiten luz a diferente longitud de onda, y un soporte informático que puede producir una imagen computerizada de los resultados de la hibridación.

30 25. Método según la reivindicación 2, que comprende:

a) extraer el ácido nucleico de una muestra sospechosa de contener dichos genomas memoria minoritarios en una cuasiespecie viral;

35 b) amplificar al menos un fragmento de ácido nucleico de dicha cuasiespecie viral;

c) clonar los fragmentos de ADN amplificados en la etapa b) en un vector adecuado;

40 d) determinar la secuencia del genoma mayoritario de la cuasiespecie viral para el fragmento amplificado;

e) amplificar los fragmentos de ADN clonados en la etapa c) y marcar los fragmentos amplificados con un compuesto marcador;

45 f) poner en contacto, en una reacción de hibridación, los fragmentos amplificados y marcados de la etapa e) con los fragmentos amplificados directamente del ácido nucleico de la cuasiespecie viral de la etapa b); y

g) resolver las diferentes secuencias virales e identificar las mutaciones indicativas de los genomas memoria minoritarios presentes en la cuasiespecie viral.

50 26. Método según la reivindicación 25, en el que el clonaje de los fragmentos de ADN amplificados en la etapa b) se realiza en un plásmido de alto número de copias.

55 27. Método según la reivindicación 25, en el que el compuesto marcador utilizado en la etapa e) para marcar los fragmentos de ADN amplificados y clonados en la etapa c) es ³²P.

28. Método según la reivindicación 25, en el que la resolución de las diferentes secuencias virales, etapa g), se realiza mediante:

60 g.i) fraccionamiento de los híbridos formados en la etapa f) mediante electroforesis en gel de poliacrilamida en condiciones no desnaturalizantes;

g.ii) identificación de la existencia de genomas memoria minoritarios por el número de mutaciones en función a un salto repentino en la movilidad electroforética;

65 g.iii) extracción de los ADN hibridados y fraccionados en el gel de poliacrilamida por elución;

g.iv) amplificación de los fragmentos eluidos en la etapa g.iii);

ES 2 273 835 T3

g.v) secuenciación de los fragmentos amplificados en la etapa g.iv); y

g.vi) comparación de las secuencias deducidas en la etapa g.v) e identificación de las mutaciones indicativas de los genomas memoria minoritarios presentes en la cuasiespecie viral.

5

29. Método según la reivindicación 2, que comprende:

a) extraer el ácido nucleico de una muestra sospechosa de contener dichos genomas memoria minoritarios presentes en la cuasiespecie viral;

10

b) amplificar al menos un fragmento de ácido nucleico de dicha cuasiespecie viral;

c) determinar la secuencia del genoma mayoritario de la cuasiespecie viral para el fragmento amplificado;

15

d) opcionalmente, clonar el fragmento de ácido nucleico amplificado en un vector;

e) secuenciar el fragmento clonado; y

20

f) comparar las secuencias deducidas en las etapas c) y e), e identificar las mutaciones indicativas de los genomas memoria minoritarios presentes en la cuasiespecie viral.

25

30

35

40

45

50

55

60

65

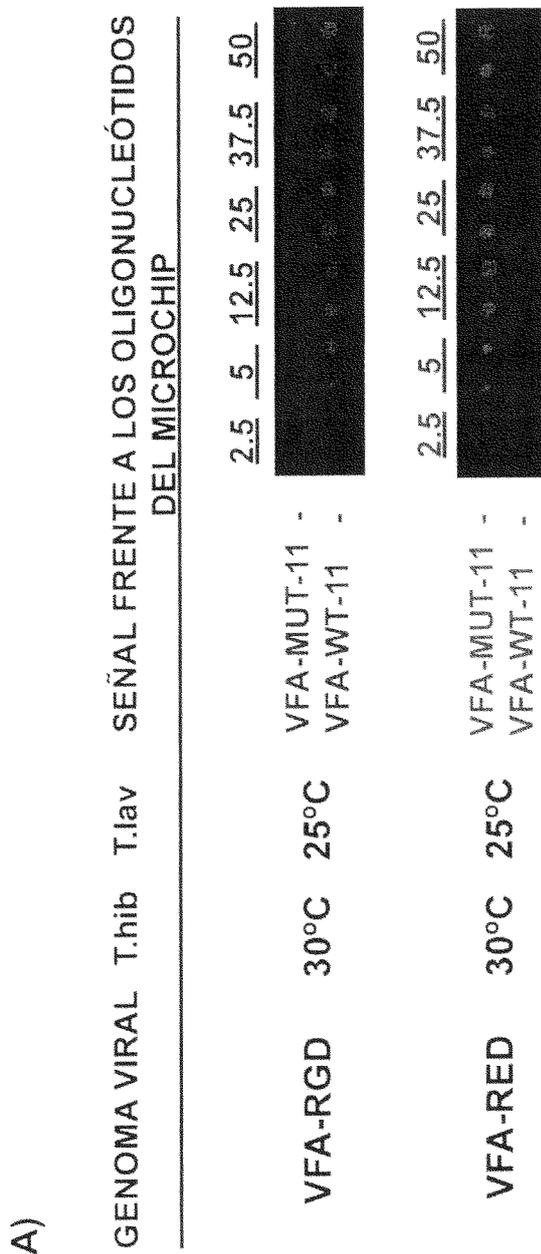
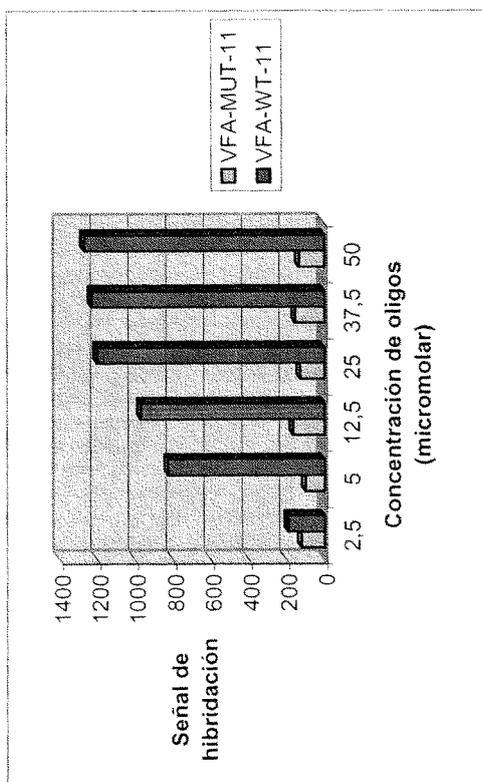


Figura 1 (a)

B)

CUANTIFICACIÓN DE LAS SEÑALES DE HIBRIDACIÓN

VFA-RGD



VFA-RED

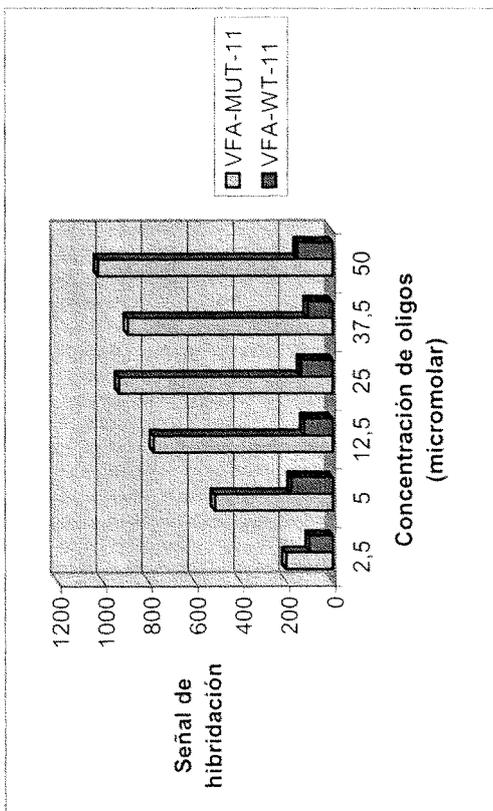


Figura 1 (b)

A)

Hibridación del genoma VFA-RGD con los puntos (duplicados) que contienen mezclas de los oligonucleótidos VFA-WT-11/VFA-MUT-11. Sobre los puntos se indica el % de VFA-WT-11 en la mezcla:

100 99.9 99 95 90 80 70 50 40 30 20 10 5 1 0.1 0



Hibridación del genoma VFA-RED con esos mismos puntos (duplicados) que contienen mezclas de los oligonucleótidos VFA-WT-11/VFA-MUT-11. Sobre los puntos se indica en este caso el % de VFA-MUT-11 en la mezcla:

0 0.1 1 5 10 20 30 50 60 70 80 90 95 99 99.9 100



Figura 2 (a)

B)
CUANTIFICACIÓN DE LAS SEÑALES DE HIBRIDACIÓN

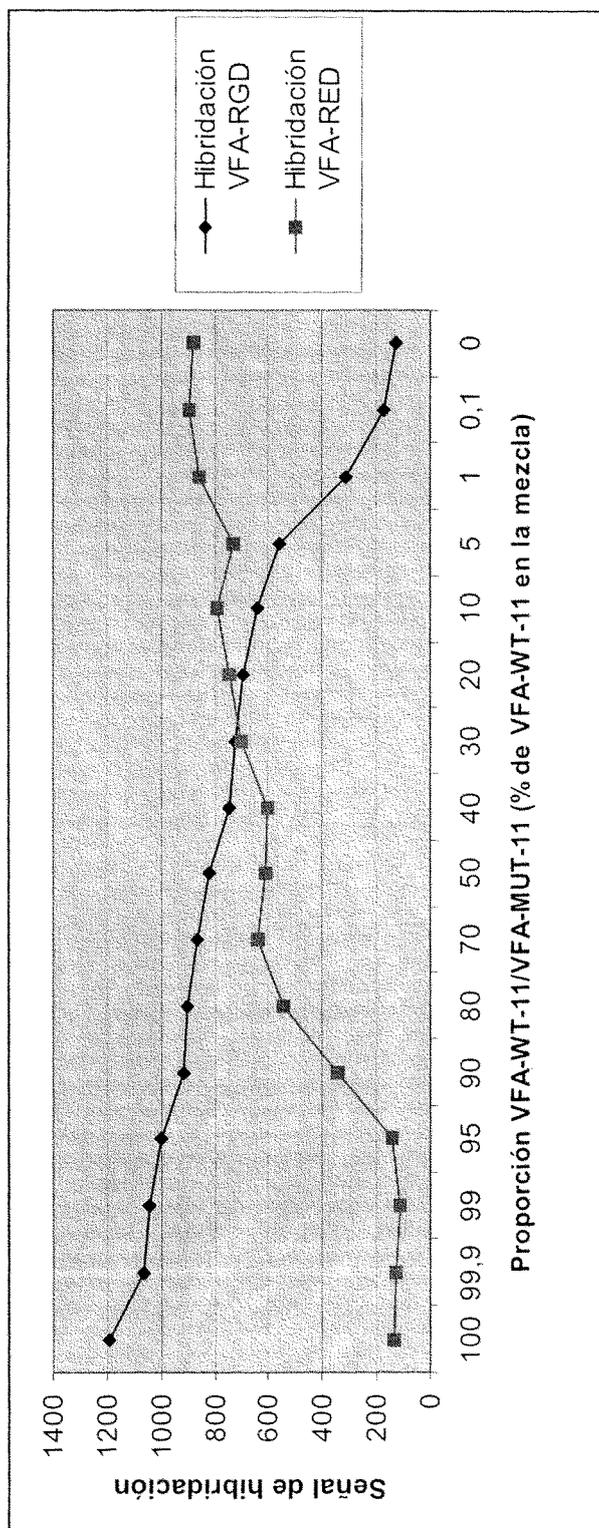


Figura 2(b)

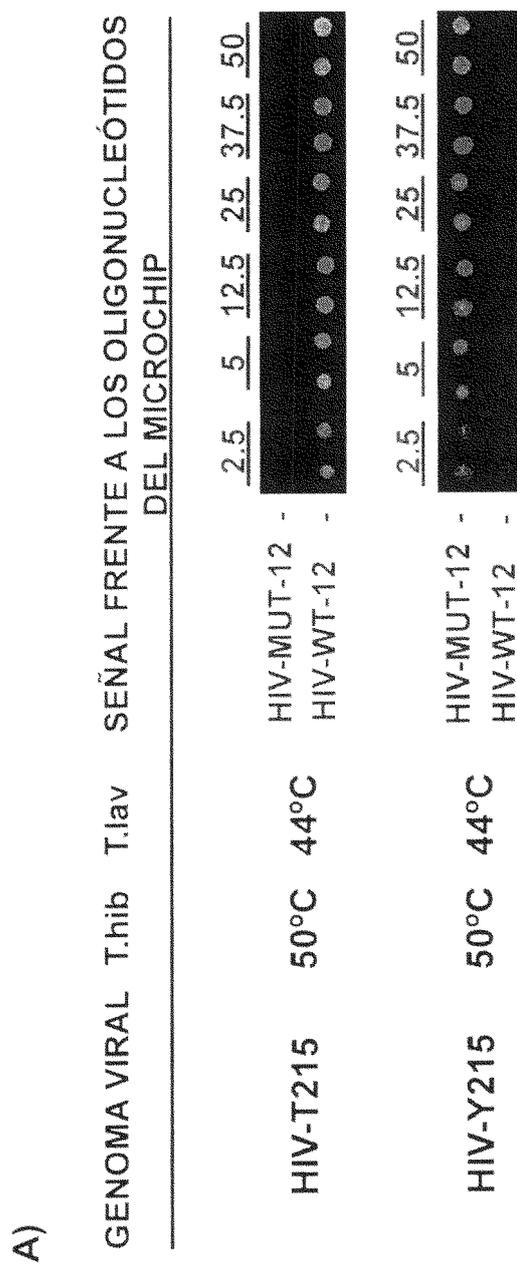
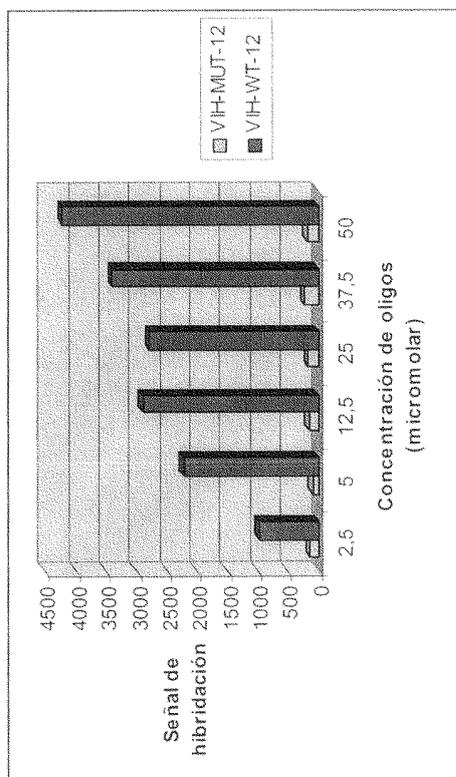


Figura 3 (a)

B)

CUANTIFICACIÓN DE LAS SEÑALES DE HIBRIDACIÓN

VIH-T215



VIH-Y215

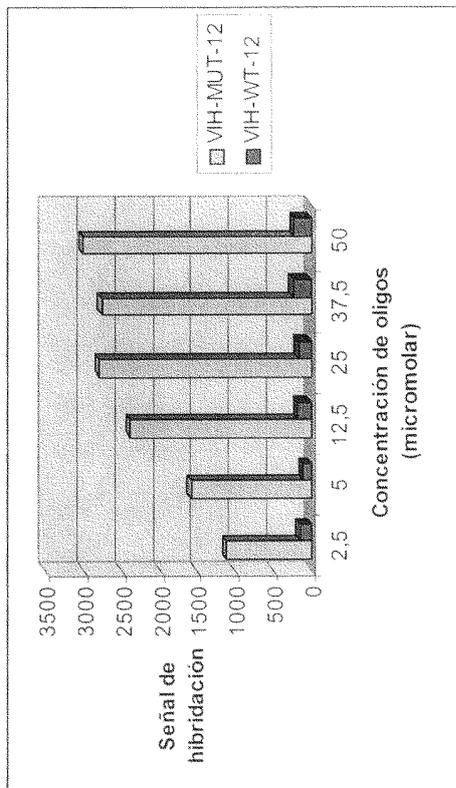


Figura 3(b)

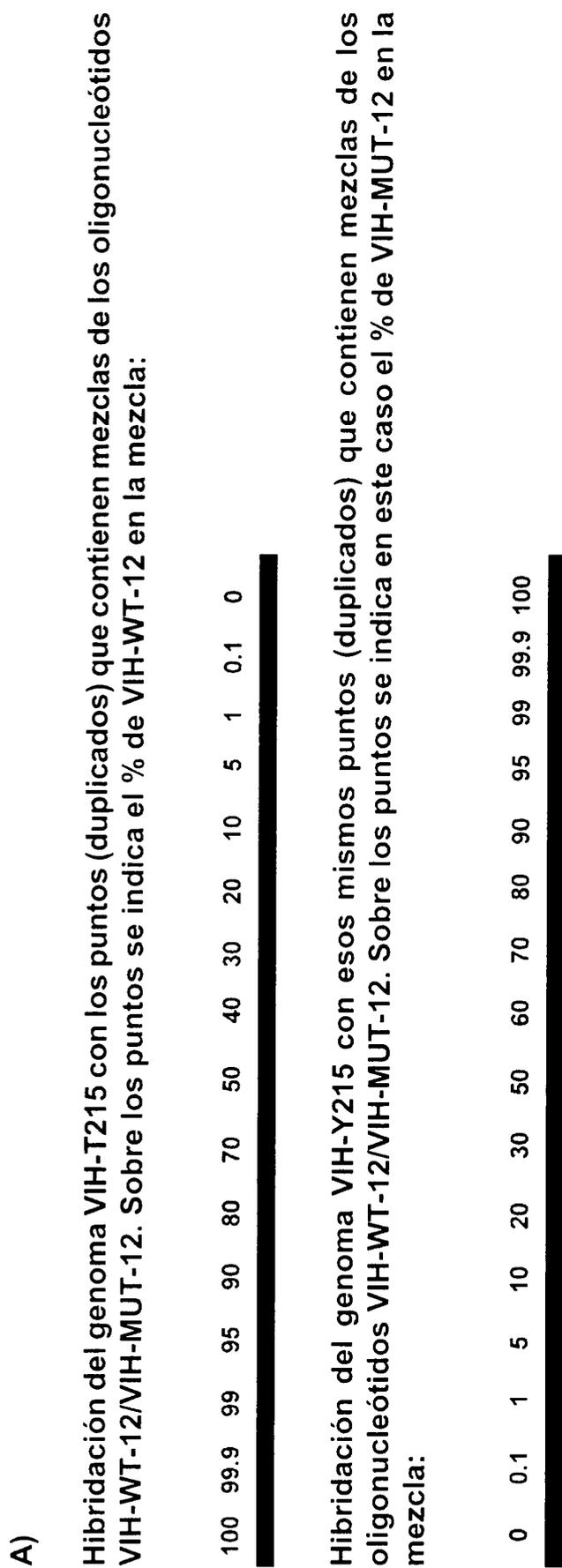


Figura 4 (a)

B)

CUANTIFICACIÓN DE LAS SEÑALES DE HIBRIDACIÓN

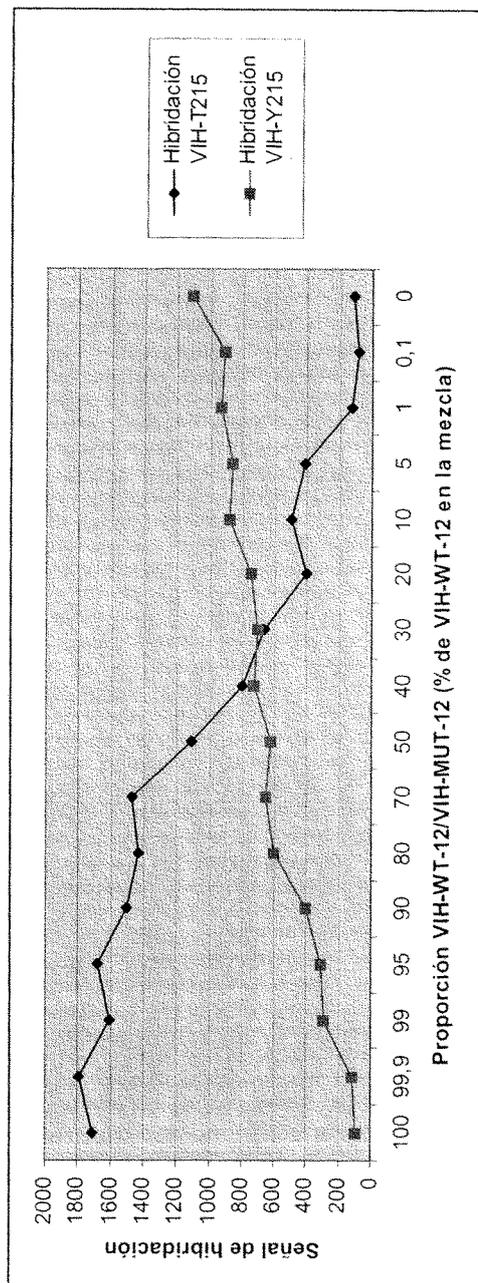


Figura 4 (b)

A) Alineamiento de secuencias de los genotipos 1-a y 1-b del VHC implicadas en respuesta a interferón

Región 6967-7086

```

1-a  CATCTCTCAAGGCAACTTGCACCGCTAACCATGACTCCCCTGATGCTGAG
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1-b  CTTCTTTGAAGGCGACATGCACTACCCACCATGACTCCCCGGACGCTGAC

CTCATAGAGGCCAACCTCCTATGGAGGCAGGAGATGGGCGGCAACATCAC
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
CTCATCGAGGCCAACCTCTTGTGGCGGCAGGAGATGGGCGGAAACATCAC

CAGGGTTGAGTCAGAAAACA
  | | | | | | | | | |
TCGCGTGGAGTCAGAGAATA
    
```

Región 2319-2351

```

CCGAGCTCAGCCCGTTACTGCTGACCACTACA
| | | | | | | | | | | | | | | | | |
CAGAACTCAGCCCGCTGCTGCTGTCTACAACA
    
```

B) Secuencias del VHC relacionadas con respuesta a INF y actividad ribozima

<u>Posiciones en el genoma</u>	<u>Secuencia</u> (genotipo 1-b)
175-215 ¹	TGCCAGGACGACCGGTCCTTTCTTGGATCAACCCGCTCAA
310-350 ¹	AGTGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCACG
2319-2351 ²	CAGAACTCAGCCCGCTGCTGCTGTCTACAACA
6967-7086 ²	CTTCTTTGAAGGCGACATGCACTACCCACCATGACTCCCCGGACGCTGAC CTCATCGAGGCCAACCTCTTGTGGCGGCAGGAGATGGGCGGAAACATCAC TCGCGTGGAGTCAGAGAATA

¹ secuencias relacionadas con ribozimas, ² secuencias relacionadas con respuesta a INF.