

VARIABILIDAD GENÉTICA DEL VIRUS DE LA HEPATITIS C

María del Carmen Maroto Vela y Federico García García

Departamento y Servicio de Microbiología. Facultad de Medicina y Hospital Universitario San Cecilio. Granada.

El virus de la hepatitis C (VHC) fue inicialmente reconocido como agente productor de la hepatitis no-A, no-B, y clonado posteriormente en 1989. Se calcula que un 3% de la población mundial, aproximadamente entre 170 y 200 millones de personas, está infectada, correspondiendo 4 millones a USA y unos 5 millones a Europa. La prevalencia varía según los distintos países, su grado de desarrollo o, incluso, las diferentes zonas y circunstancias sanitarias dentro del mismo país. En España se considera que la población infectada oscila entre el 1-2%. Alrededor del 80% de los infectados evolucionan a la cronicidad, entre un 10-20% desarrollan una cirrosis en un plazo de unos 20 años y, anualmente, el 2% de los pacientes desarrollan un hepatocarcinoma.

ESTRUCTURA GENÓMICA DEL VHC

Únicamente se comentará algunas de sus características con la finalidad de poder entender claramente el problema de la variabilidad genética. Es un virus RNA, de polaridad positiva, con un genoma de 9,5 Kb, y un tamaño que oscila entre 55 y 65 nm. Tiene una cápside proteica y una envuelta, y taxonómicamente se encuentra encuadrado en los Flavivirus, aunque hoy día se considera un nuevo género, *Hepacivirus*. Como luego veremos, presenta una serie de genotipos, subtipos y cuasiespecies.

El genoma (figura 1) contiene un marco abierto de lectura único de, aproximadamente, 3000 aminoácidos, flanqueado por regiones no traducidas altamente conservadas, denominadas 5' y 3' UTR. De las dos, la región 5' es la mejor conservada, la que menos varía, con analogías superiores al 98%, y cuya principal función es permitir la unión del ribosoma de las células hospedadoras al RNA vírico en la estructura conocida como IRES (*internal ribosome entry sites*).

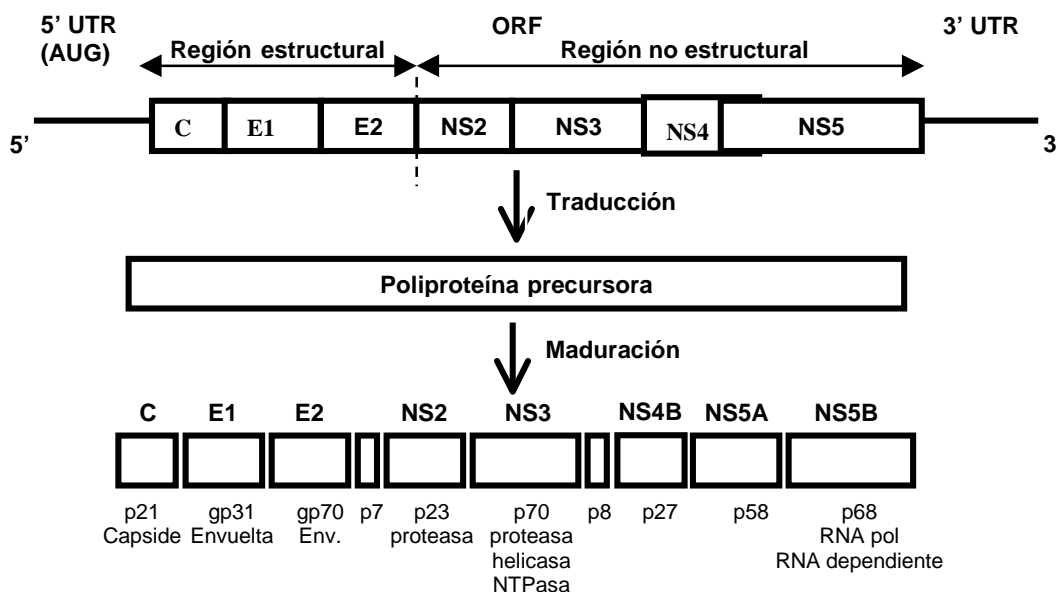


Figura 1. Representación esquemática del genoma del VHC.

El marco de lectura presenta dos regiones: una estructural y otra no estructural. La primera es capaz de codificar las proteínas de la cápside (C) y las gp31 y gp70 (E1 y E2) de la envuelta. Entre las regiones E1 y E2 se encuentra la zona denominada HVR1 (hipervariable) que permite al virus su escape del sistema inmunitario y, por lo tanto, su capacidad de influencia en la aparición de infecciones persistentes y de fracasos terapéuticos. La segunda región, no estructural, codifica para toda una serie de enzimas con acción proteasa, helicasa, RNA-polimerasa dependiente de RNA, etc. Dentro de esa región, es importante reseñar el papel de NS3 y, sobre todo, NS5 por presentar ésta el sitio de unión a la PKR (protein-kinasa) y la zona ISDR (región determinante de la sensibilidad al interferón), ambas implicadas en los fenómenos de variabilidad y resistencia al tratamiento.

VARIABILIDAD GENÉTICA DEL VHC

De entre los virus de interés en patología humana, el virus de la hepatitis C es uno de los virus con un mayor grado de diversidad genética que se han estudiado hasta el momento. La heterogeneidad genética que presenta puede ser intragenoma, dando lugar a las **cuasiespecies** víricas, e intergenómica, que da lugar a los **genotipos** y **subtipos**.

La elevada cinética de replicación viral y la baja fidelidad de la enzima responsable de la replicación (la RNA polimerasa dependiente de RNA) son los dos principales factores que explican la elevada variabilidad genética de este virus. El virus tiene una vida media de 2,5 h en sangre y existe una alta producción diaria de partículas virales (10^{12}) en los pacientes con infección crónica; la cinética de replicación viral es, por lo tanto, superior incluso a la del virus de la inmunodeficiencia humana (VIH). En segundo lugar, la enzima que se encarga de la replicación tiene una tasa de error aproximada de 10^{-4} . Desgraciadamente no existen modelos experimentales para el cultivo del virus, lo que dificulta la obtención de datos aunque, siguiendo el modelo propuesto para el VIH, la probabilidad de una mutación puntual será del orden de 10^{-4} y de una mutación doble de 10^{-11} , lo que se traducirá en la producción diaria de, aproximadamente, 3.300 virus distintos al virus parental. De este modo, será fácil comprender que la población que infecta a un individuo es una mezcla muy heterogénea de genomas muy relacionados entre sí, con una homología superior al 98%, y que se denominan **cuasiespecies**, siendo responsables de la variabilidad intragenoma.

Recientemente, un equipo de investigadores ha estimado que el virus parental del que se originaron todas las variantes que hoy existen apareció hace, al menos, 2000 años. Si aplicamos los datos de cinética de replicación, de tasa de error de la enzima y de probabilidades de mutación, podremos entender como en todos estos años se han ido seleccionando variantes tan diversas que alcanzan un grado de diferencias en el ácido nucleico suficiente como para hablar no ya de variantes dentro de un mismo genoma, sino de variaciones intergenoma. En efecto, un modelo experimental en chimpancés (tomando muestras en la fase aguda de la infección experimental y en la fase crónica, 8 años y medio después) ha demostrado que la tasa de mutación (secuenciando genomas completos en ambas muestras) es de aproximadamente $1,44 \cdot 10^{-3}$ sustituciones por año, datos que apoyan la hipótesis anterior de que la diversidad genética existente en el VHC es el resultado de la acumulación de mutaciones durante los últimos dos milenios.

La variabilidad intergenoma da lugar a los conceptos de **genotipo**, **subtipo** y **aislado**. Se denominan **genotipos** a aquellos genomas cuyo grado de homología se encuentra entre el 66-69%; se designan con un número arábigo y, hasta el momento, se han descrito 6 genotipos mayores y hasta 11 genotipos distintos. Dentro de un mismo genotipo, cuando el grado de homología se encuentra entre el 77-80% se habla de **subtipo**; se designan con una letra, que seguirá al número que nombra al genotipo; hasta la fecha se han descrito más de 100 subtipos distintos. Dentro de un mismo subtipo, se denomina aislado a aquellos genomas en los que el grado de homología no es superior al 91-95%. La figura 2 muestra los porcentajes de homología y la nomenclatura de las variantes del VHC.

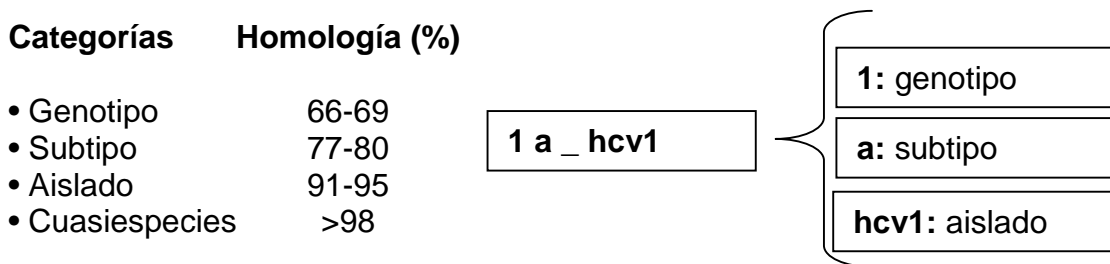


Figura 2. Homología y nomenclatura de las variantes del VHC.

Existen pacientes infectados por más de un genotipo o subtipo, lo que se denomina infección mixta. En un mismo paciente sólo se han descrito por el momento variantes intragenoma, y no variantes intergenoma, lo que quiere decir que no se ha demostrado el cambio de genotipo o de subtipo. Según esto, si tenemos en cuenta la tasa de mutación anual del VHC, podemos deducir que para que en un mismo hospedador se pudiera originar el cambio de subtipo desde el virus parental deberían pasar al menos 50-60 años. Todo esto puede cambiar en el futuro, si se confirman los datos recientes que describen el primer caso de recombinación intergenotipo en el VHC.

TÉCNICAS PARA LA DETECCIÓN DE GENOTIPOS

Las técnicas que permiten establecer la variante de VHC se pueden clasificar en moleculares, que determinan el genotipo, y serológicas, que determinan el serotipo. Para determinar el genotipo del VHC la técnica de referencia es la secuenciación de ácidos nucleicos y el posterior análisis filogenético de la secuencia obtenida. Se pueden estudiar varias regiones del genoma de VHC (NS5b, E1, *core*) ya que aquí es donde se pueden localizar con mayor exactitud las diferencias que caracterizan a los distintos genotipos y subtipos; sin embargo, la elevada variabilidad genética que caracteriza a esta región del genoma condiciona que los protocolos de amplificación para la posterior secuenciación sean extremadamente complejos, incluyendo la amplificación por separado de cada uno de los genotipos. Por lo tanto, la secuenciación de esta zona del genoma está, por el momento, restringida a determinados laboratorios de investigación especializada en el VHC. Por esto, para utilizar las técnicas moleculares para la detección de genotipos, generalmente recurrimos a la amplificación de una parte del genoma mucho más conservada, la región 5' no codificante o región 5'UTR. Además, la utilización de esta región como base para la determinación del genotipo presenta una ventaja adicional: nos sirve el amplificado que hemos obtenido para la determinación de RNA-VHC (cualitativa o cuantitativa) que estemos usando de rutina en el laboratorio ya que la mayoría de estos ensayos amplifican dicha región.

Las técnicas moleculares comerciales que podemos emplear en el laboratorio son básicamente dos: INNO-LiPA® HCV II y Truprep® HCV 5'NC Genotyping Kit (ambos de Bayer Diagnostics). El primero es un ensayo basado en la hibridación inversa: una vez obtenido el amplificado de la región 5'UTR, este se desnaturaliza, y se enfrenta a una tira de nitrocelulosa que contiene sondas que permiten la hibridación del producto de amplificación. Dichas sondas son específicas de genotipo/subtipo. El producto de la hibridación se somete posteriormente a un revelado de tipo inmunoenzimático. Todo el proceso queda automatizado, y finalmente se interpreta la tira comparándola con un patrón que proporciona la casa comercial. El procedimiento y los patrones de interpretación se resumen en la figura 3.

Truprep® HCV 5'NC Genotyping Kit se basa en la secuenciación bidireccional del producto de amplificación de la región 5'UTR. Para ello necesitamos purificar el producto de amplificación, y secuenciarlo mediante el empleo de cebadores marcados con distintos fluoróforos (síntesis dirección 3') y (síntesis dirección 5'), y dideoxi-terminadores (ddATP, ddCTP, ddGTP y ddTTP). Para cada muestra se requieren 4 reacciones de secuenciación para determinar en cada una la

secuencia de A, C, G y T. Al final un *software* de interpretación alinea las secuencias con las de referencia y establece un informe que indica el genotipo, subtipo y aislado con el que muestra una mayor homología. En esta técnica se secuencia un fragmento de 183 bases, de las que 66 son responsables de las diferencias entre genotipos/subtipos/aislados. En el resto de las posiciones no existen diferencias entre los genotipos ya que, como sabemos, la región que se secuencia es altamente conservada. Esta técnica es de mayor complejidad técnica que la citada con anterioridad.

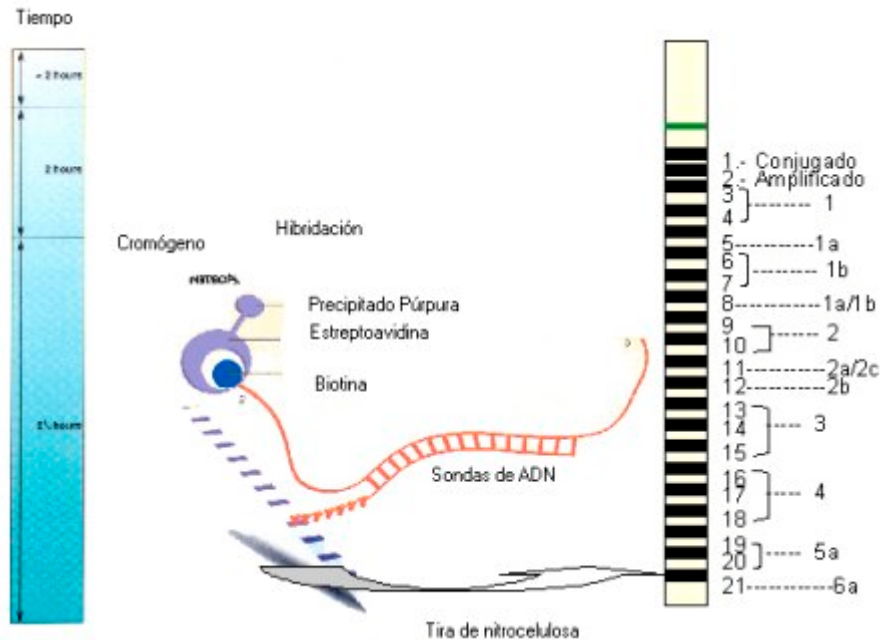


Figura 3. Fundamento e interpretación del INNO- LiPA® HCV II.

Ambas técnicas comerciales muestran una muy buena correlación para la determinación del genotipo y subtipo de VHC, pero hay que destacar que ninguna muestra una eficacia superior al 80% para la estimación del subtipo. Aunque de menor difusión, la PCR específica de tipo y el estudio del polimorfismo de los fragmentos de restricción del producto de amplificación (PCR-RFLP), que hasta hace poco se consideraba como la técnica de referencia, también se pueden utilizar para la determinación de genotipos. Actualmente se usan en pocos laboratorios debido a su menor facilidad para incorporarlas al trabajo habitual del laboratorio de diagnóstico.

Como alternativa a los métodos moleculares se han propuesto diversos sistemas para la determinación del serotipo infectante. Estos sistemas no se han impuesto, en gran parte debido a la pobre especificidad para la discriminación de la mayoría de los subtipos, a la falta de sensibilidad (más del 20 % de las muestras no son tipables), y a problemas de reactividad cruzada entre genotipos, lo que origina un gran número de resultados discordantes. Además, mediante estas técnicas se detectan con mayor frecuencia infecciones mixtas.

Mediante métodos moleculares podemos estudiar las cuasiespecies virales. Para ello, no existen técnicas comerciales y los métodos de laboratorio son laboriosos y complejos. La mejor aproximación a este estudio consiste en amplificar la región hipervariable (E2) y clonar el producto de PCR, para posteriormente secuenciar un número suficiente de clones (generalmente se acepta el 10%) y establecer el número de secuencias diferentes, que corresponderán al número de cuasiespecies (mayoritarias) presentes en la muestra. Puesto que esta técnica es extremadamente compleja, se recurre a métodos de relativa menor complejidad, que se basan en la amplificación de la región hipervariable para posteriormente evaluar el número de cuasiespecies presentes en la muestra mediante SSCP (*single stranded conformation polymorphism*) o HA (*heteroduplex analysis*).

IMPLICACIONES DE LA VARIABILIDAD GENÉTICA DEL VHC

En la tabla 1 aparecen reflejadas las diferentes implicaciones de la variabilidad del VHC en la patogenicidad, en el diagnóstico, en los fenómenos de resistencia al tratamiento, así como en la epidemiología del este virus.

Tabla 1. Implicaciones de la variabilidad genética del VHC.

Sobre la patogenicidad
<i>En la gravedad de la infección</i>
<i>En las manifestaciones extrahepáticas</i>
<i>En la aparición de hepatocarcinoma</i>
En el diagnóstico
<i>Métodos serológicos</i>
<i>Métodos moleculares</i>
En la resistencia al tratamiento
<i>Papel de NS5: PKR y región ISDR</i>
<i>Papel de NS3</i>
Sobre la epidemiología
<i>Variabilidad geográfica</i>
<i>Marcadores epidemiológicos</i>
<i>Vacunaciones</i>

Importancia en la patogenicidad

- En la gravedad de la infección.* Un hecho hasta hace muy poco tiempo aceptado era que las infecciones por los genotipos 1b evolucionaban más fácilmente hacia la cirrosis y el hepatocarcinoma. Recientemente, en la Conferencia de Consenso de los *National Institutes of Health* norteamericanos de junio de 2002 se cuestiona si existe realmente esa correlación, por lo que no deberá utilizarse como marcador virológico absoluto. Se ha demostrado que la región HVR1 del VHC contiene ciertos epítomos que podrían seleccionar mutantes capaces de escapar al sistema inmune. En general, podemos decir que existe un alto número de cuasiespecies en función de la gravedad del proceso y que las hepatitis agudas resolutivas tienen un número de cuasiespecies menor que las crónicas activas. Por otra parte, se ha comprobado que en el genotipo 1b es más frecuente encontrar hepatitis B oculta (ADN positiva, HBsAg negativa) y con una peor respuesta al interferón. Igualmente estaría relacionado con hepatitis fulminante.
- En las manifestaciones extrahepáticas.* Existen numerosos estudios que asocian la infección por el VHC con la crioglobulinemia, aún cuando no es bien conocida su relación con los diferentes genotipos. En cualquier caso, una alta variabilidad podría ser responsable de ciertas manifestaciones autoinmunes como la ya citada, la presencia de autoanticuerpos, etc.
- En la aparición de hepatocarcinoma (HCC).* El gen de la PKR (proteína mayor inducida por el interferón), se considera un supresor de tumores, controlador de la homeostasis y del crecimiento celular. En ciertas condiciones la interacción de PKR con la zona NS5 del virus lleva consigo la inactivación de aquélla, lo que facilitaría el desarrollo del hepatocarcinoma. Por otra parte, la expresión de otra región del VHC en determinadas líneas celulares, la NS3, ha inducido tumores en algunos modelos experimentales. Actualmente, se están estudiando, las mutaciones surgidas en diferentes aminoácidos codificados por el genoma vírico (distintas según los pacientes tengan HCC o no), e incluso el papel de la proteína F (producto del *frame-shift* de algunas cepas 1b en la zona del *core* y cerca del codón 11).

Importancia en el diagnóstico

- a) *En el diagnóstico serológico:* la mayoría de las técnicas utilizadas como cribado se fabricaban a partir de proteínas del genotipo 1, lo que pudo conducir a una falta de sensibilidad. De hecho se demostró que la mayoría de las muestras con resultados indeterminados mediante ensayos suplementarios de segunda generación correspondían a genotipos 2, 3, 4 y 5, por lo que el diagnóstico en los pacientes con estos genotipos puede ser menos eficiente, sobre todo en áreas en las que los genotipos más prevalentes estén filogenéticamente distanciados del prototipo 1a y se empleen ensayos de segunda generación. De todos modos, estos problemas se han solucionado casi en su totalidad con los ensayos de tercera generación (tanto de cribado como suplementarios), que son los más utilizados en la actualidad.
- b) *En los métodos moleculares:* hay que destacar que los métodos de detección de RNA deben emplear cebadores dirigidos a la amplificación de la región 5'UTR; en caso contrario, se pueden dejar de detectar variantes virales dada la elevada heterogeneidad de otras regiones del genoma. Asimismo, los ensayos de cuantificación vírica pueden verse influenciados por los genotipos. Las versiones iniciales de los ensayos de cuantificación más ampliamente utilizados (Roche Amplicor® y Versant® bDNA Bayer) mostraban diferente eficacia según el genotipo. En la actualidad, dichos problemas están también prácticamente solucionados.

Importancia en el tratamiento

Éste es, sin duda, uno de los apartados en los que existe mayor acuerdo en la importancia que el genotipo ejerce. La indicación, la duración y la dosis de ribavirina (RBV) a emplear en el tratamiento combinado con interferón pegilado (Peg-IFN) se deciden según el genotipo. En la actualidad se acepta que se debe ofrecer el tratamiento combinado (Peg-IFN+RBV), en ausencia de contraindicaciones, a todos los pacientes infectados con genotipo 2 ó 3 ya que la posibilidad de obtener una respuesta virológica sostenida es superior al 70%-80%, y es suficiente tratar durante 24 semanas y con una dosis de 800 mg/día de RBV. Sin embargo, en los pacientes infectados por el genotipo 1 (y probablemente el 4 y los demás), las posibilidades de obtener una respuesta sostenida es del 40-45%, el tratamiento ha de prolongarse hasta 48 semanas y la dosis de RBV debe ajustarse entre 1000 mg/d y 1200 mg/d. En estos pacientes, la indicación del tratamiento debe ser sopesada conjuntamente con los datos de la biopsia hepática, ya que dependiendo del grado de fibrosis y de la actividad necroinflamatoria puede plantearse retrasar el tratamiento. Mientras que la respuesta al tratamiento está claramente ligada al genotipo, el papel de los subtipos es menos claro. Este fenómeno se podría explicar por diferentes mecanismos.

- a) *Vía NS5.* La PKR es una proteína que, además de ser supresora de tumores, es capaz también de bloquear la expresión de los genes víricos. La interacción de la PKR con la región NS5-A puede inhibir la autofosforilización de dicha enzima y, secundariamente, su actividad kinasa. Existe una interacción entre la proteína NS5-A y los genotipos 1a y 1b y el sitio catalítico de PKR, siendo éste el primer mecanismo descrito por el que el virus podría hacerse resistente al interferón. Por otra parte, ya Enamoto publicó el incremento de mutaciones en una región NS5-A, denominada ISDR (región de sensibilidad al interferón), preferentemente situada entre los aminoácidos 2209-2248 en los aislados de individuos 1b (malos respondedores), y que el estudio de la ISDR podría permitir predecir el tipo de respuesta. En una publicación reciente, se confirma que el principal foco de mutaciones relacionadas con la respuesta al interferón se encuentra entre las regiones 2350-2390. En cualquiera de los casos, creemos importante reseñar que por el simple recuento de mutaciones es difícil establecer una relación completa con la respuesta, porque la observación de acúmulos de mutaciones varía según el número de regiones del genoma a estudio.
- b) *Vía NS3.* Recientemente, estudiando las regiones NS3 (con acción ATPasa/helicasa) de individuos infectados, un grupo de investigadores ha encontrado tres tipos de clones procedentes del genotipo 2a (buenos respondedores) y de los 1a y 1b (no respondedores). Pues bien, aunque la diferencia de secuencias entre todos los genotipos era muy pequeña,

alrededor del 15%, existía una variación clara en el residuo 450 que condicionaba la respuesta al tratamiento. En el caso de los respondedores, el aminoácido era una isoleucina, y en el de los no respondedores, una treonina.

Por último, se ha propuesto por parte de algunos grupos la existencia de una relación entre el número de cuasiespecies virales y la respuesta al tratamiento. De este modo, y según estos autores, el número de cuasiespecies estaría inversamente relacionado con la respuesta al tratamiento, independientemente del genotipo y de la carga viral basal. Actualmente es muy difícil asegurar que el tratamiento seleccione las cuasiespecies virales resistentes al antivírico.

Importancia en epidemiología

a) *Variabilidad geográfica.* La historia de la epidemiología del VHC está basada en la tasa de cambios de sus secuencias nucleotídicas. Según ciertos autores, a partir de modelos matemáticos, el primer ancestro común del VHC se remontaría a más de 2000 años, la divergencia de genotipos a 500 años, y las diferencias entre 1a y 1b a 300 años.

En este momento existe una gran divergencia geográfica que se explicaría por los movimientos poblacionales, el uso de drogas por vía parenteral y la contaminación por transfusiones sanguíneas. Los genotipos más repartidos son el 1, 2 y 3, responsables de la mayoría de las hepatitis C en Europa occidental, USA y Japón. El genotipo 4 es más frecuente en África del Norte, Central y Oriente Próximo. El genotipo 5 predomina en África del Sur, y del 6 al 11 en el sudeste asiático. Se debe hacer hincapié en la importancia de la variabilidad de los genotipos según los grupos de riesgo, siendo por ejemplo más frecuentes los genotipos 1a y 3a en los usuarios de drogas por vía parenteral. La figura 5 muestra la distribución geográfica actual de los principales genotipos y subtipos del VHC.

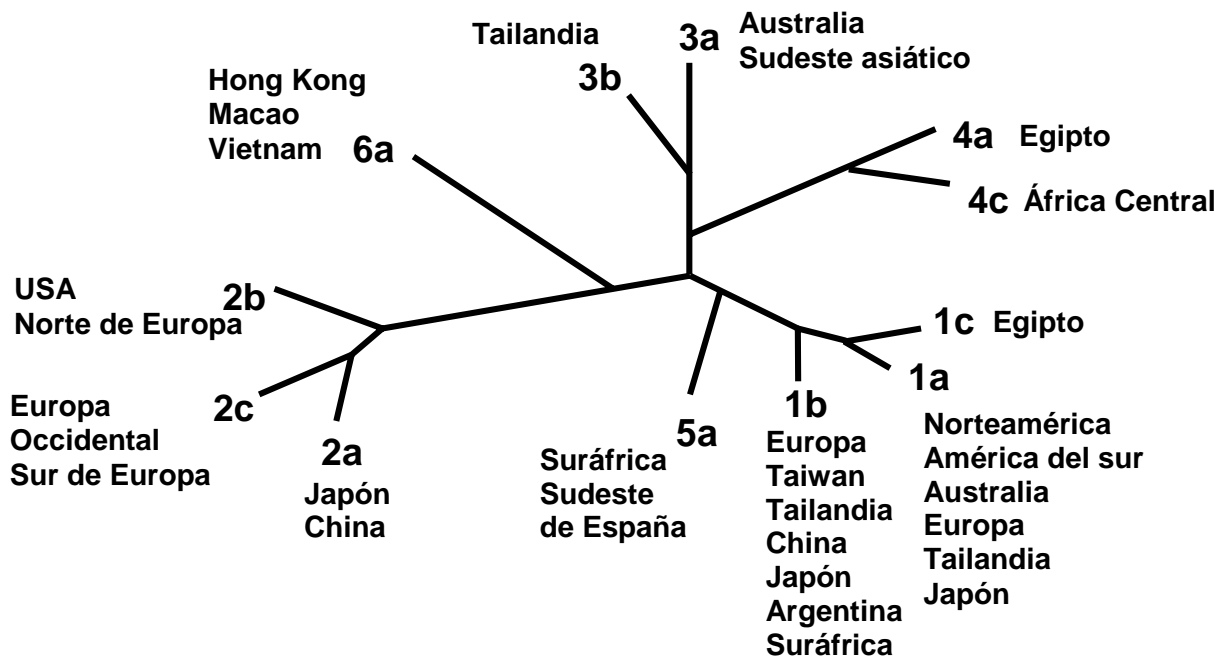


Figura 5. Distribución geográfica actual de los principales genotipos y subtipos de VHC

b) *Los genotipos como trazadores epidemiológicos.* El estudio de los genotipos puede ser útil para el control de migraciones o el comercio de productos sanguíneos, pero en ambos casos este tipo de estudios es poco determinante para analizar la transmisión entre individuos o para comprobar una contaminación por la misma cepa. Para eso son necesarios estudios filogenéticos más profundos a partir de diferentes regiones (HVR1-E2), lo que nos permitiría conocer la transmisión nosocomial, sexual o maternofamiliar.

- c) *Vacunaciones*. La gran variabilidad del VHC hace muy difícil la puesta a punto de una vacuna útil. Según diferentes estudios hechos en chimpancés, existen anticuerpos neutralizantes sólo específicos de cepa, por lo que sería necesario hacer “sopas” de antígenos que pertenecieran a diferentes cepas. Otra dificultad añadida sería la ausencia de un modelo animal (fuera del chimpancé), así como de sistemas eficaces de cultivo. No obstante, existen diferentes líneas de investigación abiertas utilizando proteínas de distintas zonas del *core* y vehiculizadas incluso por otros virus.

BIBLIOGRAFÍA

KATO N. Molecular virology of hepatitis C virus. *Acta Med Okayama* 2001; 55:133-159.

LUNEL F, STUYVER L, BRECHOT C, MAERTENS G. Mise au point sur le virus de l'hépatite C: variabilité et implications. *Transfus Clin Biol* 1998; 5:147-165.

MAJID AM, GRETCH DR. Current and future hepatitis C virus diagnostic testing: problems and advancements. *Microbes and Infection* 2002; 4:1227–1236.

ZEIN NN. Clinical significance of hepatitis C genotypes. *Clin Microbiol Rev* 2000; 13:223–235.