



Tesis de Doctorado en Ciencias Biológicas

POBLACIONES DE CUASIESPECIES DEL VIRUS DE HEPATITIS C: EVOLUCIÓN, ADAPTACIÓN Y DIÁLOGO VIRUS-HOSPEDERO.

Autor: Msc. Ricardo Recarey

Tutor: Dr. Juan Cristina Gheraldi

Laboratorio de Virología Molecular

Centro de Investigaciones Nucleares

Universidad de la República

Montevideo, Uruguay

2015

Resumen

El virus de la hepatitis C infecta a más de 170 millones de personas en el mundo. Mientras que un tercio de los pacientes infectados tienen una enfermedad hepática activa y progresiva que conduce a la cirrosis y a hepatocarcinomas, otros dos tercios tienen una enfermedad leve, no progresiva.

Una mayor comprensión de la dinámica evolutiva intrapaciente de las poblaciones virales de VHC y su relación con virulencia, resistencia a drogas y terapias anti-VHC, así como la comprensión de los mecanismos de adaptación al huésped es fundamental para facilitar el desarrollo de medidas terapéuticas más efectivas.

En la presente tesis estudiaremos esa dinámica evolutiva en diferentes contextos.

El estudio de la dinámica evolutiva durante el tratamiento en función al tipo de respuesta operada por el paciente, mostró que cada paciente presenta diferentes dinámicas de población independiente al resultado de la terapia, si bien se observan patrones evolutivos diferentes en diferentes pacientes, esto es claramente visualizado mediante el estudio de redes complejas. Las redes filogenéticas podrían pues desempeñar un importante papel en la reconstrucción de la historia evolutiva.

El estudio de la dinámica evolutiva durante un evento de transmisión materno-fetal, así como el estudio en mayor detalle hecho por pirosecuenciación, permitió destacar la importancia de los genomas minoritarios o genomas memoria dentro de la población.

La importancia de la región hipervariable, que debe de variar en secuencia para escapar al sistema inmune pero debe de conservar su estructura para reconocer los receptores celulares.

Establecemos la posibilidad de que el antigenoma de HCV pueda codificar MiRNAs capaces de regular la expresión génica del hospedador. Determinando al menos un MiRNA cuyos blancos se hallan involucrados en vías asociadas a la respuesta inmune.

INDICE

1. Introducción	6
1.1 <i>La enfermedad</i>	6
1.1.1. <i>Epidemiología del HCV</i>	6
1.1.2. <i>Patogénesis del HCV</i>	9
1.1.3. <i>Respuesta inmune</i>	12
1.1.4. <i>Transmisión del HCV</i>	16
1.1.5. <i>Diagnostico de la infección por HCV</i>	16
1.1.6. <i>Tratamiento de la infección por HCV</i>	17
1.2. <i>El Virus de la Hepatitis C</i>	22
1.2.1. <i>Clasificación</i>	22
1.2.2 <i>Estructura y organización del genoma</i>	24
1.2.2.1 <i>Regiones no codificantes (NCR)</i>	27
1.2.2.2. <i>Proteínas estructurales</i>	29
1.2.2.3. <i>Proteínas no estructurales</i>	30
1.2.2.4. <i>Marco alternativo de lectura (proteína F)</i>	33
1.3. <i>Ciclo replicativo</i>	34
1.4. <i>Variabilidad Genética del HCV</i>	37
1.4.1. <i>Genotipos</i>	37
1.4.2. <i>Distribución geográfica de los diferentes genotipos</i>	40
1.5. <i>Evolución del HCV</i>	42
1.5.1. <i>Mutación</i>	42
1.5.2. <i>Recombinación</i>	44
1.5.3. <i>Complementación</i>	45
1.6. <i>Las cuasiespecies</i>	45
1.6.1. <i>Teoría de las cuasiespecies</i>	45
1.6.2. <i>Transmisión y divergencia de las cuasiespecies</i>	48
1.6.3. <i>Compartimentación de las cuasiespecies</i>	48
1.6.4. <i>Métodos de análisis de cuasiespecies</i>	49
1.6.5. <i>Genomas memoria</i>	51
1.7. <i>Representación de historias evolutivas</i>	53
1.7.1. <i>Arboles filogenéticos vs redes de haplotipos</i>	53
1.7.2. <i>Tipos de Redes</i>	54

1.7.2.1. Arboles Filogenéticos.....	56
1.7.2.2. Las redes de división.....	56
1.7.2.3. Las redes reticuladas.....	56
2. Materiales y métodos.....	58
2.1. Construcción de los Datasets.....	58
2.2. Alineación de las secuencias.....	59
2.3. Análisis de uso de codones.....	59
2.4. Análisis coalescente bayesiano de Cadenas Monte Carlo de Markov.....	59
2.5. Análisis filogenético.....	60
2.6. Análisis de redes.....	60
2.7. Identificación de epítopes.....	62
2.8. Mapeo de sustituciones aminoacídicas.....	62
3. Objetivos.....	63
3.1. Objetivo general.....	63
3.2. Objetivos específicos.....	63
4. Objetivo específico 1.....	63
4.1. Resultados del objetivo específico 1.....	65
4.1.1. Análisis de uso de codones.....	65
4.1.2. Análisis evolutivo durante el tratamiento.....	66
4.1.3. Análisis evolutivo antes, durante y posterior al tratamiento.....	67
4.1.4. Análisis coalescente bayesiano.....	68
4.2. Discusión del objetivo específico 1.....	84
4.2.1. Análisis de uso de codones.....	84
4.2.2. Análisis evolutivo durante el tratamiento.....	84
4.2.3. Análisis coalescente bayesiano.....	86
5. Objetivo específico 2.....	86
5.1. Resultados del objetivo específico 2.....	88
5.1.1. Análisis Bayesianos de la NS5A y la HVR1/E2.....	88
5.1.2. Análisis de diversidad media dentro de las sub-poblaciones.....	89
5.2. Discusión del objetivo específico 2.....	93
5.2.1. Análisis de diversidad media.....	93
6. Objetivo específico 3.....	93
6.1. Resultados del objetivo específico 3.....	97
6.1.1. Análisis filogenéticos y de redes de los haplotipos.....	97

6.1.2. Predicción de epítopes.....	98
6.2. Discusión del objetivo específico 3.....	105
6.2.1. Análisis filogenético y redes de los haplotipos.....	105
6.2.2. Predicción de epítopes.....	105
6.2.3. Análisis de las mutaciones asociadas a los epítopes.....	106
6.2.4. Mapeo de la sustitución en el aa 440.....	106
7. Objetivo específico 4.....	108
7.1. Resultados del objetivo específico 4.....	110
7.1.1. Análisis coalescente bayesiano.....	110
7.1.2. Análisis del árbol filogenético de la HVRI.....	110
7.1.3. Análisis de redes de mediana.....	111
7.1.4. Co-evolución de sitios de aminoácidos en la HVRI.....	111
7.2. Discusión del objetivo específico 4.....	118
7.2.1. Análisis coalescente bayesiano.....	118
7.2.2. Análisis del árbol filogenético de la HVRI.....	119
7.2.3. Análisis de redes de mediana.....	120
7.2.4. Co-evolución de sitios de aminoácidos en la HVRI.....	120
8. Objetivo general B.....	121
8.1. Análisis de miRNAs asociados al HCV.....	121
8.2. Materiales y Métodos.....	122
8.2.1. Análisis de secuencias.....	122
8.2.2. Confirmación de las secuencias de los posibles pre-miRNA....	124
8.2.3. Similitud nucleotídica de Pre-miRNA con miRNAs humanos....	124
8.2.4. Identificación de secuencias de miRNA maduros.....	124
8.2.5. Conservación de los miARN maduros.....	125
8.2.6. Predicción de estructura secundaria de pre miRNA.....	125
8.2.7. Hibridación entre virales pre-miRNAs y miRNAs humanos.....	125
8.2.8. Predicción de posibles objetivos y análisis funcional.....	125
8.2.9. Hybridization entre los blancos y los miRNA maduros.....	126
8.3. Resultados.....	126
8.3.1. Predicción de posibles pre-miRNAs virales en el antigenoma...	126
8.3.2. Predicción de miRNAs maduros.....	129
8.3.3. Conservación de las secuencias de miRNAs maduros.....	129

8.3.4. Predicción de blancos para el MiRNA predicho	130
8.3.5. Hibridación eficaz de miARN de la cepa H77 de HCV.....	130
8.4. Discusión.....	142
9. Conclusiones.....	146
10. Perspectivas.....	148
11. Agradecimientos.....	148
12 Bibliografía.....	149

Abreviaturas

AAD.- antivirales de acción directa.
BOC.- Boceprevir
ETR.- Pacientes respondedores al final del tratamiento.
HCC.- Carcinoma hepatocelular.
hsa-miR.- MicroRNA humanos.
HVR1.- Región hipervariable 1.
INF α /RBV.- Interferón α en combinación con Ribavirina.
IRES.- región interna de entrada ribosomal.
ISDR.- región sensible al interferón.
MCMC.- Cadenas Monte Carlo de Harkov.
MFE.- Mínimo de energía libre.
MJ.- Red Mediana de Uniones
nAbs.- anticuerpos neutralizantes.
NR.- Pacientes no respondedores.
pre-miRNAs.- Precursores de microRNAs
RBV.- Ribavirina.
RSCU.- Uso Relativo de Codones Sinónimos.
RVS.- Pacientes con respuesta viral sostenida.
TVR.- Telaprevir
vsRNAs.- miRNAs codificados por virus.

1. Introducción

1.1 La enfermedad

El término hepatitis significa inflamación del hígado, y puede ser causada por diversos mecanismos, incluyendo agentes infecciosos, siendo los virus los principales causantes. Las hepatitis virales son causadas principalmente por al menos cinco virus; llamados virus de la hepatitis A, B, C, D y E. Particularmente, los tipos B y C desarrollan a una enfermedad crónica en cientos de millones de personas, son la mayor causa de cirrosis y cáncer de hígado en el mundo (WHO, 2003).

El virus de la hepatitis C (HCV) fue el primer virus que se descubrió por medio de técnicas moleculares. Aunque la hepatitis C, como enfermedad, fue reconocida por primera vez en 1975, como una hepatitis no A no B (Feinstone *et al.* 1975), todos los medios disponibles para identificar el agente causante de la enfermedad fallaron hasta 1989, cuando el grupo de Choo y colaboradores clonaron y secuenciaron un fragmento del genoma del HCV (Choo *et al.* 1989).

El uso de métodos moleculares ha sido un instrumento fundamental en el descubrimiento del virus así como en los avances realizados en el estudio de la hepatitis C. No obstante, a pesar del extraordinario progreso acontecido en las últimas décadas, el HCV continúa siendo uno de los mayores desafíos para virólogos y médicos.

1.1.1. Epidemiología del HCV

El HCV es el principal agente causante de hepatitis esporádicas no A no B y de hepatitis post-transfusionales adquiridas parenteralmente en todo el mundo (Kuo *et al.* 1989). El último reporte de la Organización Mundial de la Salud (OMS) sobre la prevalencia del virus estima que aproximadamente el 3% de la población mundial se encuentra infectada con el virus, afectando a 170 millones de personas. Además, se estima que cada año se infectan entre 3 y 4 millones de personas y mueren, por año, más de 350 mil personas (Sharma y Feld 2014). Las muertes asociadas al HCV se deben principalmente a cirrosis o carcinoma hepatocelular (HCC), con 211.000 y 155.000 muertes por año respectivamente (Perz *et al.* 2006).

La prevalencia del HCV varía según la región geográfica siendo de 5,3% para África, 4,6% para el este del Mediterráneo, 1,03% para Europa, 2,15% para el sudeste asiático y 3,9% para el oeste del Pacífico y para la región Latinoamericana del 1,3%; existiendo variaciones notorias entre los diferentes países de estas regiones (Sy y Jamal 2006). Existen 14 millones de personas crónicamente infectadas en América, 400.000 en Oceanía, 16 millones en el Este Medio, 17,5 millones en Europa, 28 millones en África, 83 millones en Asia y 8 millones para la región Latinoamericana (Lavanchy 2011), variando entre los diferentes países y entre regiones de un mismo país (Szabo *et al.*, 2012; Méndez-Sánchez *et al.*, 2010) (Figura 1).

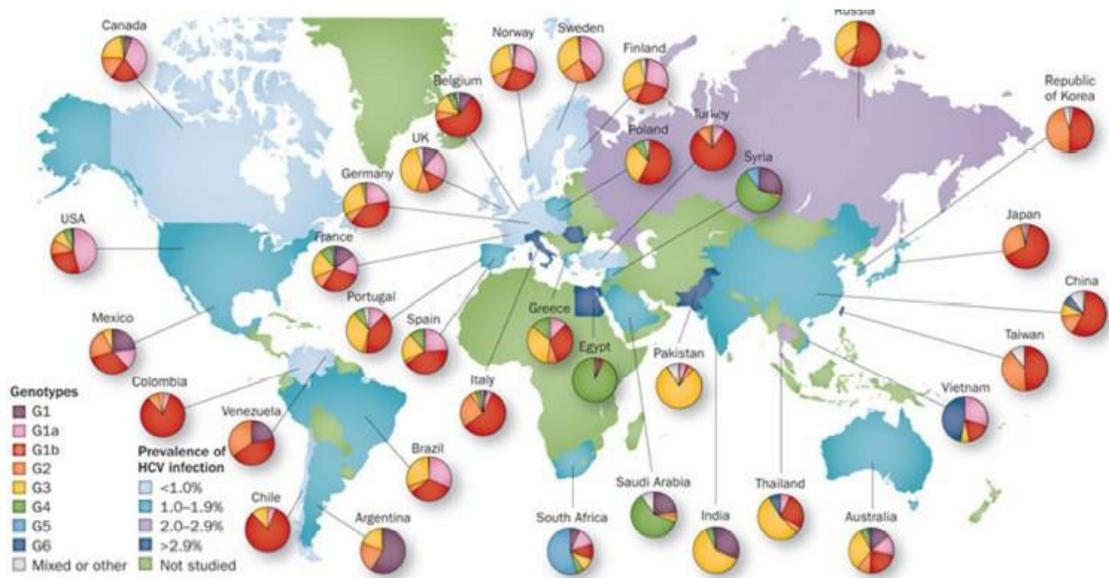


Figura 1. Prevalencia del HCV y distribución de los genotipos en el mundo. Se muestra la prevalencia de la infección del HCV en el mundo, así como también los genotipos circulantes, en los países que se ha estudiado (Hajarizadeh *et al.* 2013).

1.1.2. Patogénesis del HCV

El HCV infecta preferentemente células hepáticas, pero a pesar de ser este el sitio principal de replicación viral, existen reservorios extrahepáticos tales como los linfocitos, las células epiteliales del intestino y las células del sistema nervioso central.

La partícula viral posee una baja densidad de flotación lo que le permite interactuar con las lipoproteínas del suero. Esta interacción permite suponer que al menos una porción de los virus circulantes pueden ser internalizados en lipoproteínas del hospedero para formar una partícula “lipoviral”, lo que le permitiría obtener un camuflaje que le permitiría escapar a los anticuerpos neutralizantes.

La mayoría de los pacientes con infección aguda no padecen síntomas clínicos aparentes, aunque pueden aparecer síntomas típicos de una hepatitis aguda como anorexia, malestar abdominal, hepatomegalia, náuseas, vómitos, fiebre, astenia o letargo, mialgia e ictericia (Orland *et al.* 2001). Durante esta fase casi todos los pacientes desarrollan una fuerte respuesta inmune humoral y celular, que en muchos casos no consigue evitar la infección pero es una de las causas del daño hepático, debido a la destrucción masiva de células hepáticas por el continuo ciclo de replicación y liberación del virus y el ataque del sistema inmune contra las células infectadas. Los niveles de aminotransferasas (ALT) en suero, son indicativos del daño a los hepatocitos, aumentan entre la segunda y la octava semana tras la exposición. La hepatitis C aguda puede ser grave y prolongada, pero raramente es fulminante (Farci *et al.* 1996). También existen factores genéticos del paciente como el gen IL28B que puede predecir la eliminación del virus así como también la respuesta al tratamiento con interferón (Sharma y Feld, 2014).

Alrededor del 15% de los pacientes eliminan el virus y la infección remite por sí sola; los síntomas duran varias semanas y desaparecen conforme caen los niveles de ALT y de RNA del HCV.

En el 85% de los casos los pacientes infectados no eliminan el virus y progresan a la infección crónica, existiendo gran incertidumbre en el pronóstico de la enfermedad ya que algunas personas infectadas se recuperan completamente; mientras que otras permanecen con viremia de HCV sin evidencia bioquímica de daño hepático; algunos parecen tener una forma estática de la hepatitis crónica caracterizada por niveles

persistentemente elevados de aminotransferasas sin síntomas manifiestos o avance de la enfermedad.

De estos pacientes el 20% puede desencadenar en cirrosis y un 1-4% desarrollan hepatocarcinoma (Rosen y Gretch, 1999; Mindikoglu *et al.* 2009) (Figura 2).

Debido a que la cirrosis puede tardar varios años en desarrollarse, la infección con HCV puede permanecer silenciosa durante décadas. Por tanto los individuos infectados, en la mayor parte de casos, suelen desconocer su estado durante largos períodos de tiempo, y la enfermedad es detectada con frecuencia en análisis de sangre rutinarios (por niveles elevados de enzimas hepáticas) o donaciones de sangre (Di Bisceglie 1998; Bostan y Mahmood 2010).

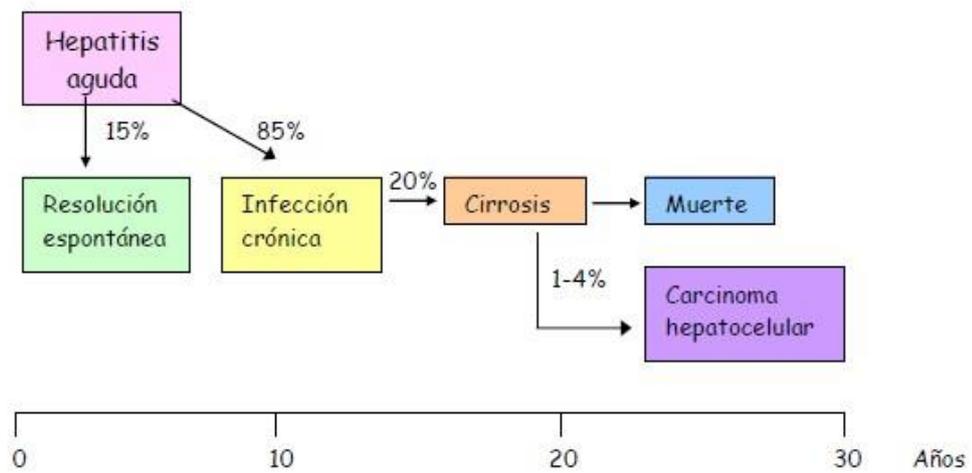


Figura 2. Esquema representando la evolución de la infección por HCV, mostrando los porcentajes esperados que desarrollen diferentes procesos. En el 85% de los casos los pacientes infectados no eliminan el virus y progresan a la infección crónica de los cuales el 20% puede desencadenar en cirrosis y un 1-4% desarrollan hepatocarcinoma.

1.1.3 Respuesta inmune

La respuesta inmune inespecífica es la primera barrera defensiva del organismo y no requiere sensibilización previa. Este tipo de respuesta es mediada por células con capacidad fagocítica y células natural killer (NK) (Chen *et al.* 1999).

La respuesta específica o adquirida se desarrolla frente a sustancias que inducen su iniciación y en ella participan prioritariamente los linfocitos y los elementos solubles liberados por los mismos. La respuesta inmune específica puede ser de dos tipos: humoral y celular. Aunque la separación de ambos tipos de respuesta es más de tipo didáctico que real, en general se considera que cuando el elemento efector final son las inmunoglobulinas formadas por los linfocitos B se trata de una respuesta tipo humoral, mientras que cuando participan los linfocitos T tanto colaboradores (Th) como citotóxicos (Tc), se trata de una respuesta tipo celular.

La respuesta inmune humoral es mediatizada por los linfocitos B, que reconocen al antígeno a través de las inmunoglobulinas de membrana. Sin embargo este estímulo no es suficiente para que se inicien los procesos de proliferación de estas células. Para ello es necesario que los linfocitos B, además del estímulo antigénico, reciban el de ciertas interleucinas (Fournillier *et al.* 2004).

Las inmunoglobulinas, al detectar al antígeno y unirse a él, actúan como transductores de la información de la presencia de los mismos, que serán posteriormente destruidos por el mecanismo más idóneo, en el que colaborarán, además del propio anticuerpo, el sistema del complemento, macrófagos, los polimorfonucleares o células K.

El término complemento engloba una gran variedad de proteínas, que interactúan en un determinado orden. Cuando se produce la activación del complemento se pone en marcha una serie de reacciones, en forma de "cascada", de tal forma que se van generando productos activos que además de influir en que la reacción prosiga tienen diferentes acciones biológicas importantes en la defensa del organismo.

La respuesta inmune de tipo celular actúa principalmente frente a bacterias y virus, así como evitando la aparición y desarrollo de células tumorales. En ella participan esencialmente los linfocitos T colaboradores y citotóxicos. Los linfocitos reconocen el antígeno mediante el receptor T (TcR) y lo hacen cuando el antígeno es degradado y procesado en el interior de las células presentadoras de antígeno (APC) y sus determinantes antigénicos son expuestos en la superficie de estas células en el seno

de una molécula del complejo principal de histocompatibilidad. Se desencadena una cascada de reacciones bioquímicas en el citoplasma de la célula T, dando así lugar al proceso de activación, proliferación y diferenciación celular. Como consecuencia de estos eventos se producirá finalmente la activación de la transcripción de los genes implicados en la síntesis de la proteína y factor implicado en una determinada función, tal como la síntesis de interleucina 2 u otros factores. La consecuencia final de este tipo de respuesta es la formación de células Th activas productoras de interleucinas y células citotóxicas (CTL) que poseen capacidad de lisar a las células que portan el antígeno que indujo su activación (Figura 3).

Las moléculas del Complejo Mayor de Histocompatibilidad (MHC) son una serie de glicoproteínas presentes en las membranas de todas las células nucleadas, entre las que se encuentran las células inmunocompetentes. Estas moléculas son esencialmente de dos tipos o clases, clase I y clase II. La separación de las funciones de los linfocitos T colaboradores CD4+ y CD8+ viene dada por el origen de los antígenos que reconocen y, en último término, por donde han sido procesados por vía exógena en el sistema endosomal de las células presentadoras de antígeno y expresados en superficie por el producto de los genes MHC de clase II. Los linfocitos citolíticos CD8+ reconocen a los antígenos que han sido procesados endógenamente en el citosol de la célula infectada y presentados en superficie por moléculas MHC de clase I (Cox A *et al.* 2005), mientras que los linfocitos CD4+ interactúan con el antígeno en el contexto de moléculas de clase II (Netski *et al.* 2007).

Las citocinas son una serie de sustancias producidas por células en respuesta a una gran variedad de estímulos y que son capaces de regular el funcionamiento de otras células. Estas sustancias pueden ser de diversos tipos entre los que se encuentran los denominados factores de crecimiento, polipéptidos que estimulan la proliferación de diferentes tipos celulares; las linfoquinas, producidas por linfocitos (sobre todo Th) y de gran importancia en la regulación del sistema inmune. La familia, globalmente denominada interferones, fueron originalmente identificadas como agentes capaces de proteger a las células frente infecciones virales.

Hoy se sabe que los interferones tienen otras muchas funciones, tales como actuar en los procesos de diferenciación y proliferación celular así como en la modulación del sistema inmunológico.

Los interferones se clasifican en dos tipos: Tipo I que comprende el IFN α , sintetizado por leucocitos, y el IFN β , sintetizado por fibroblastos, ambos producidos en

respuesta directa al virus; y el Tipo II, al que pertenece el IFN γ , sintetizado en respuesta al reconocimiento de células infectadas por los linfocitos T y células NK (Amadei *et al.* 2010) (Figura 3).

La unión de ambos tipos de interferón a sus receptores celulares desencadena una serie de procesos que conlleva la activación de la transcripción de genes (normalmente expresados a bajos niveles) que impiden la replicación del virus o interfieren con procesos virales (Thomas *et al.* 2011). A esta de serie de pasos se le conoce como la ruta JACK/STAT (JACK: Janus tirosine kinases), STAT: (señales de activación de la traducción y la transcripción). Además los interferones pueden enlentecer el proceso de crecimiento celular, haciéndolas más susceptibles a la apoptosis y limitando así la infección viral.

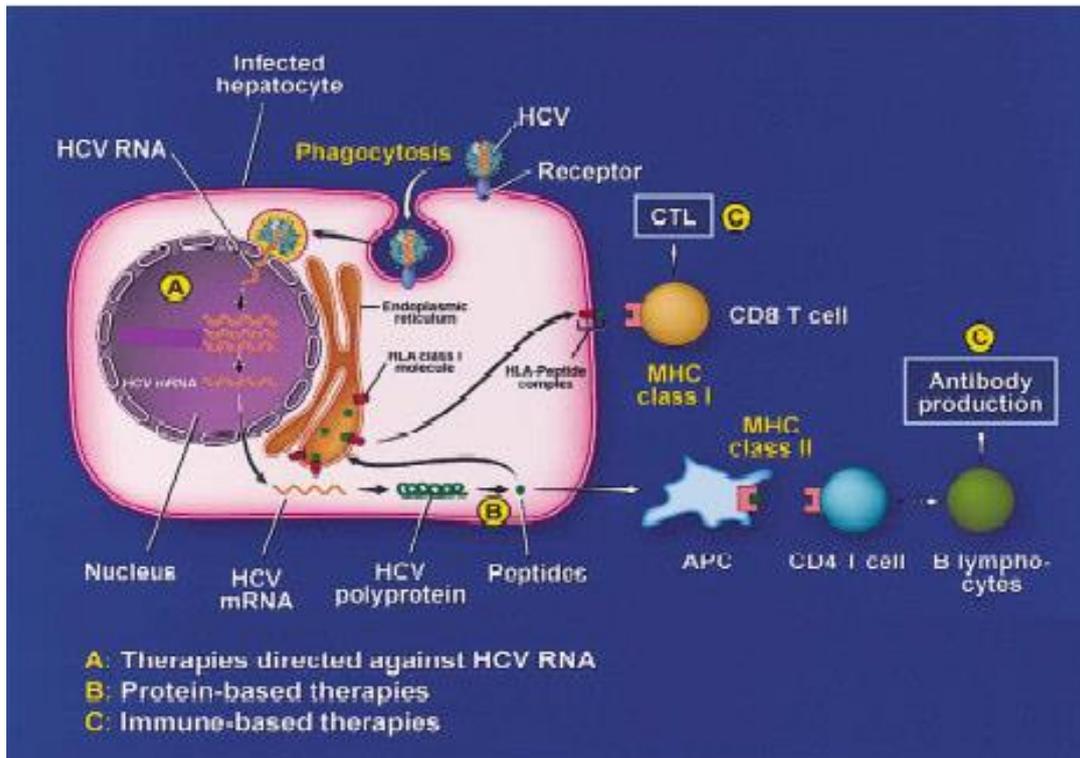


Figura 3. Respuesta humoral y celular desencadenada tras la entrada del HCV en un hepatocito (Tomado de Zein *et al.* 2002).

1.1.4 Transmisión del HCV

La principal vía de transmisión es la exposición directa o indirecta con sangre contaminada, bien sea a través de transfusiones de sangre, por consumo de drogas vía intravenosa, o por el empleo de material quirúrgico mal esterilizado.

Aunque varía según la región, depende del estado de desarrollo del país.

Los países desarrollados tienen como principal vía de transmisión el uso de drogas inyectables principalmente a través del uso compartido de jeringas, relaciones sexuales sin uso de preservativo (mayoritariamente en hombres que tienen relaciones con hombres), tatuajes o perforaciones, y también por el uso de drogas vía nasal (El-Shabrawi y Kamal, 2013; Zaltron *et al.* 2012). Actualmente, la transmisión a través de transfusiones es infrecuente debido al tamizaje por inmunoensayo enzimático y detección de ácido nucleico en la sangre utilizada y los cuidados adecuados tanto en los procedimientos como en los materiales utilizados en el área de la salud (Shepard *et al.* 2005).

Por otro lado, en los países en desarrollo, las inyecciones terapéuticas y las transfusiones de sangre sin el tamizaje adecuado tienden a ser las principales vías de transmisión, especialmente en países donde los controles en la salud son escasos o nulos, como sucede en algunas regiones de África donde las condiciones no son las adecuadas y hasta se llegan a reutilizar jeringas (El-Shabrawi y Kamal 2013; Shepard *et al.* 2005).

1.1.5 Diagnóstico de la infección por HCV

El diagnóstico virológico y el monitoreo de la infección por el HCV está basado en dos tipos de test, los métodos indirectos que identifican anticuerpos contra proteínas virales y los ensayos directos donde se identifica al virus, ya sea mediante la secuenciación del mismo o mediante la determinación de antígenos virales.

Entre la infección y la detección de una respuesta inmune específica contra HCV pueden pasar de 2 a 3 meses, lo que da lugar a la existencia de una amplia ventana inmunológica.

La presencia de RNA en sangre periférica es indicativo que el virus se está replicando activamente principalmente en el hígado. El RNA viral se detecta 1 ó 2 semanas después de la infección, en el caso de una infección aguda el RNA alcanza un

máximo y luego desaparece, en el caso de infección crónica se observa una disminución pero luego se estabiliza.

Para determinar la carga viral se pueden utilizar distintos ensayos cuantitativos, RT-PCR cuantitativa, el NASBA, el DNA Elisa de Dia Sorin o la detección de PCR en tiempo real, la OMS desarrollo el estándar WHO 96/7909 determinándose su valor en 100.000 UI/ml. Aunque no existe relación entre la carga viral y el desarrollo hacia la cronicidad.

1.1.6 Tratamiento de la infección por HCV

El objetivo de la terapia es la erradicación de la infección por HCV con el fin de evitar complicaciones hepáticas asociadas al virus y enfermedades extrahepáticas, incluyendo necroinflamación hepática, fibrosis, cirrosis, HCC y la muerte.

El punto final de la terapia es la respuesta virológica sostenida (RVS), definida por la no detección de RNA del HCV 24 semanas después de finalizar el tratamiento, según la evaluación de un método molecular sensible con un límite inferior de detección <15 UI/ml (RVS24).

La Ribavirina es un análogo de nucleósido que es incorporado al RNA viral por la RNA Polimerasa (Airaksinen *et al.* 2003; Contreras *et al.* 2002; Crotty *et al.* 2001; Severson *et al.* 2003); actuando fundamentalmente como agente mutagénico (Severson *et al.* 2003; Tam *et al.* 2001; Vignuzzi *et al.* 2005); La ribavirina también tiene propiedades inmunomoduladoras, mejorando la respuesta antiviral de las citoquinas de tipo 1, (Fang *et al.* 2000; Hong y Cameron 2002), actúa regulando en alza la expresión de genes antivirales a través de elementos de respuesta al interferón (Zhang *et al.* 2003).

Estos hallazgos sugieren que en el caso del HCV, que la ribavirina puede ejercer un efecto antiviral directo a través tanto de la mutagénesis como de la inhibición o supresión de la replicación del RNA viral.

Por otro lado, la aparición de un mutante con resistencia a ribavirina, (un único cambio de aminoácido de Phe a Tyr en la posición 415 de la polimerasa viral) en cinco individuos sometidos a monoterapia con ribavirina para la infección por HCV (Young *et al.* 2003), argumenta a favor de un efecto antiviral directo, debe tenerse en cuenta que la monoterapia con ribavirina no tiene prácticamente ningún efecto sobre la viremia y que el mayor beneficio de la ribavirina es la supresión de la recaída en lugar de la supresión de la replicación del virus.

Los interferones son glicoproteínas de la clase de las citocinas, reciben su nombre debido a su capacidad para interferir en la replicación de una amplia variedad de virus de RNA y ADN en las células hospedadoras. Se unen a receptores en la superficie de las células infectadas, activando diferentes vías de señalización en las que participan diversas proteínas antivirales (como PKR).

El interferón alfa es un tipo de interferón con 14 isoformas importantes en la respuesta antiviral. Además el INF- α -peg se encuentra unido covalentemente a polietilenglicol retardando de esta forma la eliminación del fármaco de manera que se consiguen niveles más constantes del mismo en sangre con menores dosis. Además el INF- α -peg produce simultáneamente una activación de la respuesta inmunitaria del hospedero.

El tratamiento con (INF- α) reduce efectivamente la carga viral, pero la erradicación completa del virus se logra en menos de 20% de los pacientes tratados solo con INF- α (Poynard *et al.* 1998). La ribavirina aumenta la frecuencia de la erradicación del virus en pacientes que inicialmente responden al INF- α , Pero tiene un efecto limitado en los pacientes que no responden al INF- α (Hu *et al.* 2001).

Diferentes genotipos y subtipos se correlacionan con diferente susceptibilidad al tratamiento con INF en monoterapia o con terapia combinada de INF- α /RBV). Sólo 10-20% y 40-50% de los individuos con infección crónica por el genotipo 1 del HCV en monoterapia y en terapia combinada, respectivamente, exhibir aclaramiento estable de la infección por HCV. Estas tasas son inferiores a las tasas de entre el 50 y 70 a 80% respectivamente que se observó en tratamiento de infecciones con los genotipo 2 o 3 (Pawlotsky 2003; Zeuzem 2004).

El uso de interferón- α pegilado (INF- α peg) en combinación con RBV fue un importante avance, decisivo en el tratamiento de la infección crónica por el HCV que se asocia con aumentos significativos de estos porcentajes (Fried y Hadziyannis 2004).

Sin embargo, la tasa de respuesta virológica sostenida sigue siendo insatisfactoria (McHutchinson *et al.* 1998) particularmente en pacientes infectados con el genotipo 1 (Poynard *et al.* 1998), el más prevalente en muchas regiones geográficas del mundo (Simmonds *et al.* 2005).

El tratamiento optimizado con estos agentes puede inducir una respuesta virológica sostenida en 70% a 80% de los pacientes con el genotipo 2 y 3, pero en sólo 40% a 50 % de los pacientes infectados con el genotipo 1 (Afdhal *et al.* 2011).

Las guías actuales recomiendan que los pacientes con genotipos 2 y 3, se traten con INF- α -peg y dosis bajas de RBV durante 24 semanas, lo que resulta en una respuesta virológica sostenida de aproximadamente el 80%.

Teniendo en cuenta estas altas tasas de respuesta, varios ensayos recientes han evaluado si el tratamiento más corto (12-16 semanas) podría ser rentable en estos pacientes (Petta y Craxì 2011).

El genotipo 4 ha sido considerado como difícil de tratar basado en la baja RVS, a partir de datos obtenidos basados en regímenes convencionales con INF- α -peg. El tratamiento con INF- α - peg/RBV aumenta más del 60% la tasa de respuesta virológica sostenida para este genotipo (Kamal 2011).

El genotipo 1 es tratado con INF- α -peg/RBV obteniendo baja RVS, por lo tanto la detección temprana de pacientes que no responden al tratamiento de HCV limita la exposición innecesaria a los tratamientos y sus efectos secundarios. Un algoritmo que combina anticuerpos basales contra la proteína NS4A y una PCR cuantitativa durante el tratamiento identifica pacientes no respondedores a una combinación de INF- α -peg/RBV después de 1 semana de tratamiento. La decisión de interrumpir la terapia a pacientes crónicos con genotipo 1 tratados con INF- α -peg/RBV puede realizarse con confianza después de 4 semanas de tratamiento basada en la ausencia de anticuerpos basales contra la proteína NS4A y una cantidad de RNA viral mayor de 100.000 UI/ml (Orlent *et al.* 2010).

Desafortunadamente, la terapia de INF-Peg prolongada necesaria para controlar las infecciones crónicas de HCV se asocia a menudo con efectos secundarios graves tales como fatiga, fiebre, y mialgias, por lo general, estos síntomas responden al tratamiento con agentes antiinflamatorios no esteroideos (Sharma 2010).

En algunos pacientes, el tratamiento con IFN también ha dado lugar a problemas psiquiátricos, como la depresión, la ansiedad e irritabilidad. Requiriéndose una reducción de la dosis en más de un tercio de los pacientes, y la interrupción de la administración de las drogas es necesario en aproximadamente el 10% de pacientes (Manns *et al.* 2007).

Importantes avances en el desarrollo de potentes inhibidores selectivos de la replicación del HCV se ha logrado en la última década. Algunos medicamentos recién desarrollados como telaprevir (TVR), primer inhibidor de la proteasa (IP), se espera que pueda entrar en el mercado en un futuro próximo, un gran número de medicamentos antivirales, que nos permitan elegir inhibidores selectivos para bloquear pasos

específicos de la replicación del HCV están actualmente en fase de investigación (Lok *et al.* 2010; Delang *et al.* 2010; Chatterji *et al.* 2010).

Los pacientes infectados con el genotipo 1 de HCV que no lograron erradicar el virus en un tratamiento previo con INF- α -peg/RBV deben ser considerados para el tratamiento con la triple terapia basada en inhibidores de proteasa (IP). La terapia triple aumenta las tasas de RVS de 29% a 88%, dependiendo del tipo de falta de respuesta anterior y del estado de la enfermedad hepática. Un tratamiento con INF- α -peg/RBV, sin la adición de un IP, se asocia con bajas tasas de RVS. En general, las mismas contraindicaciones se aplican a la triple terapia con Telaprevir (TVR) o Boceprevir (BOC) e INF- α -peg/RBV como a la terapia dual con INF- α -peg/RBV.

En los pacientes con cirrosis compensada, el tratamiento debe realizarse con especial cuidado, ya que la incidencia de efectos secundarios (especialmente trastornos hematológicos e infecciones graves) es significativamente mayor en la triple terapia frente a la doble terapia de INF- α -peg/RBV, sobre todo cuando la albúmina sérica es <3,5 g/dl o las plaquetas <100.000 antes de iniciar el tratamiento (EASL, 2014).

El tratamiento estándar aprobado por la Asociación Europea para el Estudio del Hígado para el cuidado y erradicación del virus en una hepatitis C crónica es la combinación de INF- α -peg/RBV y TVR o BOC para el genotipo 1 y la combinación de INF- α -peg/RBV por los genotipos 2, 3, 4, 5 y 6 (EASL, 2014).

Dentro de estas nuevas estrategias, el Sofosbuvir (SOF) es el primer compuesto para entrar en el mercado con regímenes de combinación libres de IFN.

El SOF pertenece a los inhibidores de nucleótidos de la polimerasa viral y actúa como un terminador de cadena durante el proceso de replicación del HCV, exhibiendo actividad antiviral pan-genotípica con una elevada barrera a la resistencia. Tiene una gran RVS al ser administrado junto a RBV principalmente contra el genotipo 2.

De todas formas aún se espera mejorar la RVS para el resto de los genotipos que son más difíciles de tratar, para los cuales se intentaría combinar diferentes clases de antivirales de acción directa (AAD), siendo necesarias más investigaciones sobre las interacciones droga-droga y los posibles efectos secundarios (Degasperi y Aghemo 2014). Dentro de los casos de terapia combinada de AAD libres de INF tenemos la terapia con SOF y ABT-450 que ha mostrado resultados prometedores (Zeng *et al.* 2013).

Varias publicaciones recientes han examinado de forma exhaustiva como los miRNAs celulares modulan la infección por el HCV (Pfeffer y Baumert 2010; Thibault

et al. 2013). La interacción de HCV con el miRNA del hígado, miR-122, representa un ejemplo excepcional de cómo los virus de RNA pueden usurpar la maquinaria celular del hospedero para su beneficio (Jopling, 2012). Esto condujo al desarrollo de miravirsén, un inhibidor del miR-122, por Santaris Pharma (Janssen *et al.* 2013). Miravirsén es el primer agente terapéutico basado en miRNA en entrar en ensayos clínicos en humanos para el tratamiento de enfermedades infecciosas (Lindow y Kauppinen 2012). En un ensayo de fase 2A, el tratamiento con miravirsén resultó significativo en los pacientes infectados, con perfiles de seguridad eficientes. Aunque se señaló que potencialmente podrían producirse variantes de escape a miR-122, el tratamiento con miravirsén no parece conducir a ningún escape mutacional.

En contraste con la monoterapia, el diseño y la entrega de combinaciones específicas de antagonistas e imitadores de miRNA para curar la infección por HCV o para limitar la patogénesis del hígado se han sugerido recientemente (Hoffmann *et al.* 2012).

La respuesta al tratamiento se basa en la detección del RNA del HCV y se evalúa de la siguiente manera:

Respuesta Sostenida: el RNA del HCV se negativiza durante el tratamiento y permanece indetectable por lo menos seis meses después de suspender la terapia. Es poco probable que el RNA del HCV reaparezca en la sangre de los pacientes que han logrado una respuesta sostenida.

Recidiva: los niveles de RNA del HCV se negativizan al final del tratamiento pero reaparecen una vez que se suspende el tratamiento. Rara vez, el RNA del HCV reaparece y se positiviza durante el tratamiento, y esto se denomina respuesta “penetrante” (Hadziyannis *et al.* 2004).

Sin Respuesta: el RNA del HCV permanece detectable durante el período del tratamiento. Si los niveles séricos del RNA del HCV no disminuyen significativamente entre las primeras 12 a 24 semanas después de comenzar la terapia, son pocas las probabilidades de lograr una respuesta virológica sostenida (Shiratori *et al.* 2003).

1.2. El Virus de la Hepatitis C

1.2.1. Clasificación

El HCV pertenece al género Hepacivirus dentro de la familia Flaviviridae. (Takamizawa *et al.* 1991; Cuthbert, 1994).

El nombre de la familia Flaviviridae proviene del latín flavus, que significa amarillo.

Este nombre se debe al virus de la fiebre amarilla, primer virus humano descubierto y uno de los principales miembros de la familia. Los Flaviviridae se dividen en 4 géneros llamados Pestivirus, Flavivirus, Hepacivirus y Pegivirus (Figura 4a). Algunos de los principales virus miembros del género Pestivirus son los virus de la diarrea viral bovina 1 y 2, dentro del género Flavivirus se encuentra el ya mencionado virus de la fiebre amarilla, los virus del dengue 1-5, el virus de la encefalitis japonesa, el virus del oeste del Nilo y dentro del nuevo género Pegivirus (el cual se encuentra aún en proceso de confirmación por parte de los taxonomistas) encontramos los virus GB A, C y D (Lindenbach *et al.* 2013).

Dentro del género Hepacivirus encontramos además del HCV, otros miembros que se encuentran en proceso de confirmación taxonómica como los casos del virus GB B y el hepacivirus canino; siendo este género el que posee mayor distancia genética respecto al resto de la familia (Lindenbach *et al.* 2013) (Figura 4b).

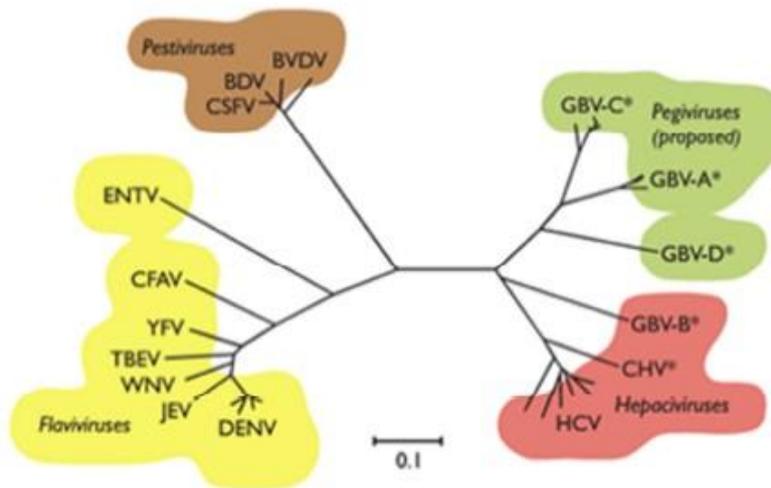


Figura 4a. Árbol filogenético de la Familia Flaviviridae realizado con el método de Neighborjoining de la polimerasa. Se distinguen los 4 géneros dentro de la familia con sus virus más relevantes. En amarillo el género Flavivirus, en marrón el género Pestivirus, en rojo el género Hepacivirus y por último en verde el nuevo género propuesto como Pegivirus (Modificado de Lindenbach *et al.* 2013).

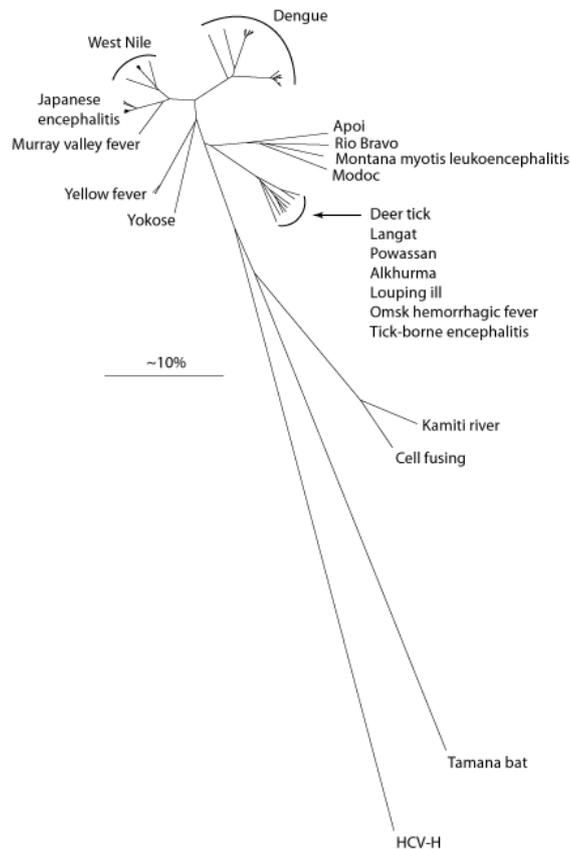


Figura 4b. Árbol Neighbor-joining de un alineamiento de genomas completos con eliminación de gaps de los flavivirus y del HCV. Se puede apreciar la distancia genética entre HCV y el resto de los flavivirus (Tomado de la base de datos LANL, www.HCV.lanl.gov).

1.2.2. Estructura y organización del genoma

EL HCV es un virus pequeño, recubierto por una envuelta, con una cadena de RNA lineal simple positiva de unas 9,6 Kb, y que posee un único marco abierto de lectura (ORF) que es traducida a una poliproteína de unos 3100 aminoácidos (Figura 5). Este ORF está flanqueado en 5' y 3' por dos regiones no codificantes: 5'NCR y 3'NCR.

Si bien el HCV tiene un único marco abierto de lectura, se ha reportado un segundo marco de lectura generado por un corrimiento en el marco a nivel del gen core (+1), el cual produce una nueva proteína denominada F (Walewski *et al.* 2001).

La poliproteína precursora es co-traduccionalmente y post-traduccionalmente procesada por proteasas celulares y virales a nivel de la membrana del retículo endoplasmático para producir 10 proteínas maduras, divididas desde el extremo 5' al 3' en 3 estructurales, core, E1 y E2, una proteína, p7, que separa las proteínas estructurales de las no estructurales y 6 proteínas no estructurales, NS2, NS3, NS4A, NS4B, NS5A y NS5B (Lindenbach *et al.* 2013) (Figura 6).

El primer corte lo realiza una proteasa del hospedador dando lugar a la única proteína de la cápside, la proteína core, a las proteínas de la envuelta E1 y E2, y a una proteína fuertemente hidrofóbica de función desconocida llamada p7. Las proteínas no estructurales restantes son procesadas por proteasas virales. El corte entre el sitio NS2 y NS3 es llevado a cabo por la proteinasa NS2-NS3 (dependiente de Zn^{+2}) de forma autocatalítica. La serinproteasa NS3 produce el corte entre NS3 y NS4A, NS4A y NS4B, NS4B y NS5A, NS5A y NS5B. El corte entre NS4B y NS5A es lento, por lo que suele encontrarse el intermediario NS4B/NS5A.

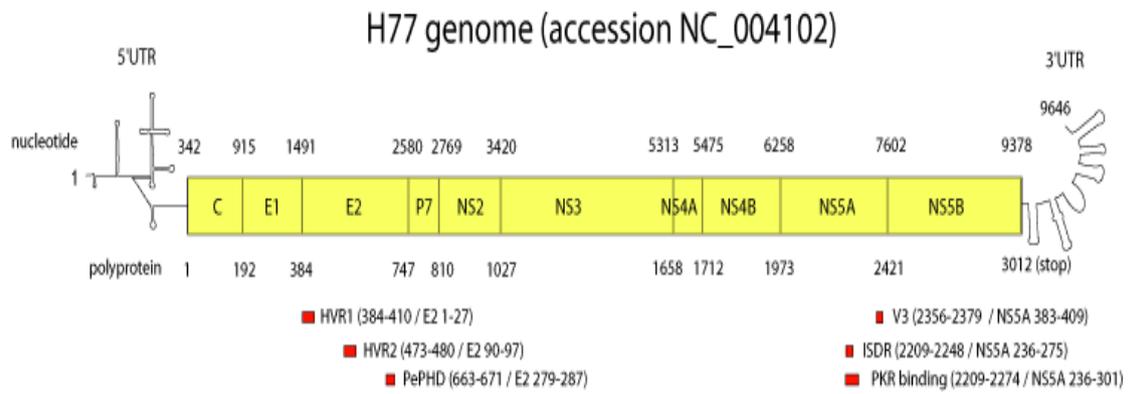


Figura 5. Esquema de la organización genómica del HCV, se trata de la cepa H77 perteneciente al genotipo 1a, se aprecian las estructuras secundarias de las regiones 3' y 5' NCR y el ORF central detallando la ubicación de las proteínas estructurales (C, E1, E2) y las no estructurales (p7, NS2-NS5B) (Tomado de Triyarni *et al* 2002).

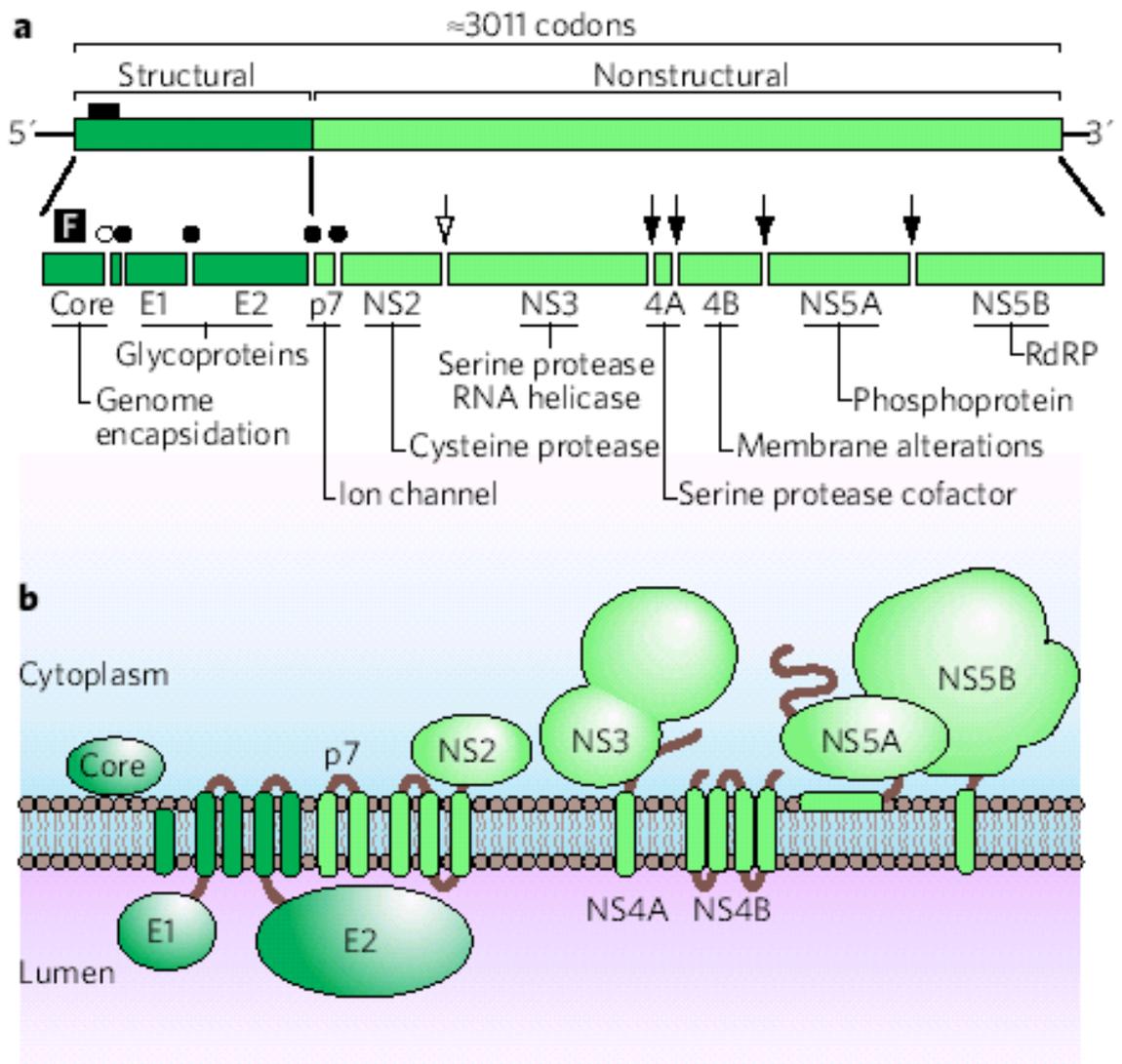


Figura 6. a) Esquema de la organización genómica del HCV: los círculos indican los sitios de clivaje dentro de la región estructural y las flechas los sitios de clivaje dentro de la región no estructural.

b) Esquema de la ubicación topológica de las proteínas del HCV con respecto a la membrana celular (tomado de Lindenbach, 2005).

1.2.2.1. Regiones no codificantes (NCR)

La 5'NCR posee una longitud de 340 nucleótidos y es la más conservada del genoma. Se compone de 4 dominios altamente conservados conteniendo una región interna de entrada ribosomal (IRES), el cual se halla formado por los dominios II, III, IV y los primeros 24-40 nt de la región codificante del core, (Tsukiyama-Kohara *et al.* 1992). Permitiendo la síntesis de la poliproteína de manera cap-independiente a través de múltiples interacciones que generan complejos binarios muy estables sin la utilización de factores de iniciación canónicos celulares (Krekulová *et al.* 2006).

El corazón del IRES que es donde se va a reclutar la subunidad ribosomal de 40S, se ubica en la parte basal del dominio III, comprendiendo los subdominios IIIa, IIIc, IIId y IIIe, los subdominios IIIabc forman las 4 puentes esenciales para unirse a la subunidad ribosomal de 40S y al factor de elongación eucariota F3 (eIF3), estructura que fue determinada por estudios de cristalografía con rayos X (Kieft *et al.* 2001) (Figura 7).

La 3' NCR de 230 nucleótidos consiste en estructuras estables tallo-bucle y un tracto interno poli (U) (Bostan y Mahmood 2010). Esta región aumenta la eficiencia de traducción del genoma viral y se cree que sirve como cebador para la síntesis de la cadena negativa del RNA (Krekulová *et al.* 2006).

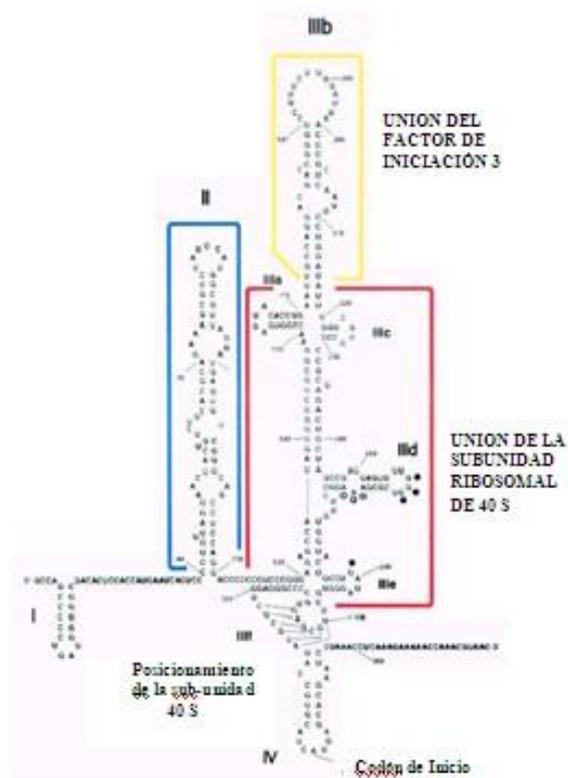


Figura 7: Esquema de la estructura de la región 5' NCR se observan los 4 dominios (I-IV); el IRES (internal ribosome entry site) está constituido por los dominios II, III y IV, unidos a los primeros 24-40 nucleótidos del core el codón de inicio AUG se halla ubicado en el dominio IV, se indican las interacciones entre el IRES y la subunidad 40S ribosomal y con el eIF3 (tomada de Sarnow 2003 con modificaciones).

1.2.2.2. Proteínas estructurales

La proteína core tiene capacidad multimerizadora e interactúa con las proteínas de la envuelta, existe en 2 o 3 formas con una masa molar de 21 kDa, 19 kDa y 16 kDa (Lo *et al.* 1996; Matsumoto *et al.* 1996). Aunque también se le atribuyen otras funciones como jugar un papel importante en la regulación del crecimiento de las células infectadas con HCV y en el desarrollo de hepatocarcinoma (Moriya *et al.* 1998).

Es una proteína básica altamente conservada, es el principal componente de la nucleocápside del HCV, forma un complejo ribonucleoproteico con el genoma viral.

La secuencia del genoma que codifica para esta proteína está altamente conservada incluso dentro de los diferentes genotipos de HCV (Krekulová *et al.* 2006), y presenta sitios antigénicos inmunodominantes, entre los aa 17-37 existe un epítipo conformacional de un motivo hélice-bucle-hélice al que recientemente se le ha adjudicado la actividad chaperonina de ésta proteína durante la dimerización del RNA (Gaël *et al.* 2004).

E1-E2 son las mayores proteínas estructurales, constituyen la envuelta y son esenciales para la penetración del virus mediante su unión al receptor y la fusión con la membrana endosomal luego de la endocitosis, ciertas moléculas proteicas como el CD81 humano y algunos glicosaminoglicanos que se hallan en la superficie de la célula del huésped participan en la fusión celular inducida por HCV (Takikawa *et al.* 2000); Los extremos N terminal de ambas proteínas indican que estas son escindidas de la poliproteína en los aa 383 y 746 respectivamente.

E1 (27 kDa) y E2 (61 kDa) son glicoproteínas transmembrana de tipo I con 6 potenciales sitios de glicosilación, poseen un ectodominio N-terminal de 160 y 334 aa respectivamente y un corto dominio transmembrana de 30 aa que contribuye al anclaje a la membrana del RE formándose heterodímeros E1E2 (Beeck *et al.* 2004).

Dentro de la proteína E2 se hallan muy variables de suma importancia tanto para la infectividad como para la integridad estructural y la función del heterodímero, en estas regiones hipervariables (HVR) la variabilidad alcanza al 80% entre los distintos genotipos, los primeros 27 aa del ectodominio de E2 constituyen la región HVR1, y entre los aa 91 a 97 de E2 se halla la otra región hipervariable (HVR2) (Penin *et al.* 2004; Mc Caffrey *et al.* 2011).

La proteína p7 se localiza en la unión entre la región estructural y la no estructural, es una pequeña proteína de membrana perteneciente a la familia de las

viporinas de 63 aa (7 kDa) que atraviesa 2 veces la membrana con sus extremos C y N terminales orientados hacia el lumen del RE lo que le permite mediar en la permeabilidad iónica de la membrana (Griffin *et al.* 2003), dado que se oligomeriza para formar una canal catiónico y parece ser de importancia en el ensamblado y maduración del virus y en consecuencia en la infectividad (Cook y Opella 2011).

1.2.2.3. Proteínas no estructurales

Hacia el extremo 3' encontramos la región que codifica para las restantes 6 proteínas no estructurales: NS2, NS3, NS4A, NS4B, NS5A y NS5B.

NS2 es una proteína hidrofóbica transmembrana no glicosilada, la mayor parte de la secuencia de NS2 se requiere para la función de la proteasa NS2-NS3 dependiente de Zn^{+2} , que es responsable del clivaje autocatalítico que la separa del resto de la poliproteína (Lindenbach 2013).

NS2 interacciona con el extremo C-terminal de la proteína celular hepática específica pro apoptótica (CIDE-B) la que pertenece a la “familia de factores CIDE inductores de apoptosis” inhibiendo su función (Erdtmann *et al.* 2003).

NS3 es una serin-proteasa (70 kDa) con funciones de helicasa involucrada en distintos cortes proteolíticos. El dominio serin-proteasa comprende los 189 aa N-terminales de la proteína que requiere asociarse a su cofactor, la proteína NS4A, este cofactor la estabiliza y la activa para realizar los clivajes 4A/4B; 4B/5A y 5A/5B (Malcolm *et al.*, 2006). Una treonina en la posición 631 es la que regula la autoproteólisis de la poliproteína (Wang *et al.* 2004). El dominio helicasa-NTPasa está constituido por los 442 aa terminales tiene varias funciones entre las que se halla el desenrollamiento de regiones de RNA con estructuras secundarias (Frick *et al.* 2004), es esencial para la replicación del RNA del HCV y también tiene un rol en el ensamblaje de partículas virales al unirse a la proteína core. Sin embargo, su función precisa viene siendo difícil de dilucidar (Morikawa *et al.* 2011; Lindenbach 2013).

NS4A–NS4B ambas proteínas son escindidas de la poliproteína viral por la serin-proteasa NS3.

NS4A es una pequeña proteína de 54 aa (8 kDa) que actúa como cofactor de NS3, inhibe la respuesta de las células Th-1 a través de la estimulación de la producción

de isoleuquina 10 (IL-10) y la inhibición de IL-12 por parte de los monocitos (Brady *et al.* 2003).

NS4B es una proteína de membrana asociada al RE, cruza 5 veces la membrana con el dominio C-terminal orientado hacia el citosol y el N-terminal se transloca hacia el lumen post-traduccionalmente. La región N-terminal comprende dos alfa hélices anfipáticas de las cuales la segunda tiene el potencial de atravesar la bicapa de la membrana. La región C-terminal comprende, aparentemente, una alfa hélice altamente conservada, una alfa hélice anfipática asociada a la membrana y dos sitios de palmitilación. NS4B interactúa con otras proteínas no estructurales virales y puede unirse al RNA viral, alberga una actividad NTPasa (Gouttenoire *et al.* 2010). Por último, NS4B tiene un papel en el ensamblaje viral al inducir una alteración específica en la membrana del RE que sirve para la fijación del complejo de replicación (Lindenbach 2013).

NS5A–NS5B son escindidas de la poliproteína por la acción de la NS3 proteasa utilizando a NS4A como cofactor, NS5A es una fosfoproteína que se le puede hallar en una forma fosforilada de 56 kDa y en una forma hiperfosforilada de 58 kDa.

Consta de 3 dominios funcionales y se halla involucrada en la replicación viral, la resistencia al interferón y la apoptosis (Liang *et al.* 2006), contiene el motivo GDD, que es común a todas las polimerasas virales, es dependiente de cationes divalentes (Mn²⁺ o Mg²⁺), a pH cercano al neutro y una muy baja concentración de sales. La tasa estimada de polimerización es de 150-200 nucleótidos por minuto y no depende de la concentración de polimerasa (Krekulová *et al.* 2006). Su función fundamental es la respuesta ante los tratamientos con interferón (INF), existe en ella una región de 40 aa denominada ISDR (Nousbaum *et al.* 2000). Dicha región mas los 26 aa siguientes constituyen el dominio de unión a PKR (proteín-kinasa) (Gale *et al.* 1998).

La interacción NS5A-NS5B es crítica para la replicación (Shimakami *et al.* 2004). Por otro lado NS5A posee numerosos motivos poliprolina (PxxP), que median la interacción con proteínas celulares conteniendo dominios Src homólogos 3 (SH3), estando relacionada esta interacción con la evasión al sistema inmune (MacDonald *et al.* 2005).

NS5B es una RNA polimerasa RNA dependiente (Hwang *et al.* 1997), se encuentra anclada al retículo endoplasmático por su hélice transmembrana C-terminal, con el dominio polimerasa en el lado citosólico.

Ningún dominio metil-transferasa se encuentra dentro de la proteína NS5B del HCV, de acuerdo con una estrategia de traducción de RNA viral que implica un sitio de entrada interno del ribosoma tanto para Hepacivirus como para Pestivirus (Caillet-Saguy *et al.* 2014).

La estructura cristalina de la NS5B del HCV fue la primera estructura completa de una RNA polimerasa dependiente de RNA. Esta estructura se asemeja a una mano derecha con tres subdominios denominados, dedos, palma y pulgar. En la estructura de NS5B de HCV además se encuentran otros tres elementos estructurales especiales.

Primero, NS5B alberga una extensión de los dedos llamada yema de los dedos que une los dominios dedos y pulgar en la parte posterior de la ApAd, dejando sólo un estrecho túnel para el acceso de nucleótidos entrantes al sitio catalítico.

Una segunda característica de la estructura es un bucle β que se inserta en el pulgar y que sobresale hacia el sitio activo. Este bucle β , también llamado bucle iniciador o bucle cebador, así como el cierre y el volumen del pulgar parecerían evitar la salida del extremo 5' de la cadena de RNA cebador en modelos de elongación ternarios de tipo polimerasa-plantilla-cebador. El bucle β también está presente en la polimerasa de Flavivirus y de Pestivirus, pero no en otras RNA polimerasa dependiente de RNA virales de estructura conocida. Este bucle β así como el extremo C-terminal son esenciales para la síntesis de RNA de novo. Ambas estructuras están dirigidas al sitio activo y constituyen la plataforma en la que el primer nucleótido debe colocarse. Una vez que se añade el primer nucleótido, estos dos dominios deben ser reubicados para proporcionar espacio suficiente para alojar el RNA de doble cadena (RNAdc). Esta reubicación implica una compleja reorganización estructural en la polimerasa y por lo tanto se piensa que es el paso más crítico durante la replicación del RNA.

Estas conversiones entre conformaciones abiertas y cerradas son importantes en el ciclo catalítico de algunas polimerasas de ácido nucleico como la transcriptasa reversa del VIH o la propia NS5B del HCV, entre otras (López-Jiménez *et al.* 2014).

La tercera sorpresa en la estructura de la polimerasa es el posicionamiento de unos 40 residuos downstream del dominio de la polimerasa que conectan este dominio y el anclaje transmembrana in vivo y que por esta razón se denomina enlazador. Este enlazador siempre se organiza como un largo bucle alrededor del pulgar que obstruye la parte frontal de la NS5B mediante el establecimiento de contactos específicos con el bucle β . Con esta ubicación, el enlazador impide la salida del extremo 3' de la cadena

de RNA molde y también contribuye a mantener el pulgar en una conformación cerrada (Caillet-Saguy *et al.* 2014).

Por último es importante mencionar la región Okamoto de la NS5B (del nucleótido 8282 al 8610 en referencia a la secuencia H77 con número de acceso AF009606), una pequeña región altamente conservada muy utilizada para realizar estudios de filogenia, filodinámica y filogeografía; por su capacidad para representar toda la diversidad genética de NS5B, la cual es a su vez también representativa de la variabilidad genética del genoma completo viral (Hraber *et al.* 2006).

1.2.2.4. Marco alternativo de lectura (proteína F)

HCV tiene un marco de lectura alternativo que se superpone con el gen de la proteína core. El marco de lectura superpuesto distingue al HCV de todos sus parientes virales conocidos, con la posible excepción del GBV-B. Este marco alternativo se expresa durante infecciones naturales por HCV y estimula las respuestas inmunitarias específicas. Al igual que varios genes esenciales en otros virus (por ejemplo, la polimerasa del VIH), este marco carece de un codón de inicio AUG dentro del marco, lo que sugiere que su expresión implica eventos inusuales a nivel de la traducción.

Los estudios *in vitro* indican que el cambio del marco por los ribosomas puede ser uno de varios procesos que pueden llevar a la traducción de la proteína alternativa.

El cambio de marco produce proteínas quiméricas que tienen segmentos codificados en el gen core unidos covalentemente a los aminoácidos codificados en este marco alternativo. Las funciones de las proteínas del marco alternativo de lectura en el ciclo replicativo del HCV no se conocen aún.

Los efectos observados de la proteína core incluyen la inducción de cáncer de hígado, la transformación de las células, y las alteraciones de la respuesta inmune.

Se considera también la posible utilidad de las proteínas alternativas en las vacunas (Branch *et al.* 2005; Cristina *et al.* 2005).

1.3. Ciclo replicativo

Reconocimiento y entrada: El primer paso en el ciclo de vida de un virus es la unión de la partícula infecciosa a la célula hospedera, que requiere la interacción entre un receptor de la superficie celular y una proteína de unión en la superficie de la partícula viral. En este caso están implicados la proteína CD81 y el receptor de lipoproteínas de baja densidad (LDLR). CD81 forma parte de la familia de las tetraspaninas, unas moléculas de la superficie celular caracterizadas por presentar tres dominios transmembrana y dos lazos extracelulares (Reed *et al.* 1995; Pileri *et al.* 1998).

E1 está involucrada en la fusión de membranas, y E2 actúa como una chaperona para E1, la cual en ausencia de E2 no se pliega correctamente (Bostan y Mahmood, 2010). Después de la unión inicial del HCV a sus receptores SRB1 y CD81, la partícula viral interacciona con las proteínas de unión CLDN1 y OCLN y finalmente entra en las células por endocitosis mediada por receptor. Posteriormente el genoma viral se libera en el citoplasma (Bartenschlager *et al.* 2013) (Figura 8).

Traducción y procesamiento de la poliproteína: Una vez dentro del citoplasma, el RNA genómico viral se traduce directamente ya que es reconocido como RNAm debido a que tiene la misma polaridad que un RNAm celular.

La traducción del RNA viral no está mediada por un mecanismo cap-dependiente sino por la iniciación interna de la traducción a través del IRES.

Dirigida por el IRES, la poliproteína se traduce en el retículo endoplasmático rugoso y se escinde co- y post-traduccionalmente en las proteínas maduras, tanto por la acción de proteasas de la célula huésped como por dos proteasas virales (Bostan y Mahmood, 2010).

Las proteínas virales, en conjunción con factores de la célula huésped, inducen la formación de un compartimento membranoso denominada red membranosa (MW) compuesta de vesículas simples, dobles y multi membranosas, así como gotas de lípidos (LD) (Bartenschlager *et al.* 2013).

Replicación del RNA: La replicación del RNA se produce en un sitio no especificado dentro de la red membranosa y a través de una copia de sentido negativo de RNA (RNA(-)) que sirve como plantilla para la producción de cantidades en exceso de RNA de sentido positivo (RNA(+)) para la progenie (Bartenschlager *et al.* 2013) (Figura 8).

La polimerasa viral cataliza la síntesis tanto de la hebra negativa como de la positiva. Secuencias en el extremo 3' se pueden plegar intramolecularmente y formar una horquilla que permite hibridizar y generar un extremo 3' que funciona como cebador y es utilizado para la elongación. Utilizando las altas concentraciones de GTP o ATP, la polimerasa puede sintetizar un cebador de RNA.

Otros factores virales y celulares son necesarios, como la helicasa NS3 que elimina estructuras estables en el molde de RNA y facilita la replicación o la fosfoproteína NS5A que está implicada en la regulación de la replicación del RNA. Además, un componente celular, denominado PTB, interactúa con las secuencias en el extremo 3'. La gliceraldehído-3-fosfato-deshidrogenasa se une a la secuencia poli (U) en el extremo 3'. Otras proteínas celulares involucradas en la replicación son las llamadas p87 y p130. Proteínas de otros virus también podrían potenciar la replicación al activar la transcripción de genes celulares (Bostan y Mahmood 2010).

Ensamblaje y liberación del virus: El ensamblado de las partículas virales se inicia probablemente en las proximidades del ER y las LD, donde se acumulan la proteína core y el RNA viral. La formación de partículas puede ser iniciada por proteínas core que interactúan con el genoma de RNA. Dicha interacción, no sólo puede promover una encapsidación selectivo del genoma de cadena positiva sino que también parece reprimir la traducción a través del IRES, lo que sugiere un potencial mecanismo para cambiar de traducción/replicación a ensamblado. La nucleocápside viral adquiere su envoltura (E1 y E2) por gemación a través de las membranas del RE y en este caso el virus es exportado a través de la vía secretora constitutiva (Bostan y Mahmood 2010; Bartenschlager *et al.* 2013) (Figura 8).

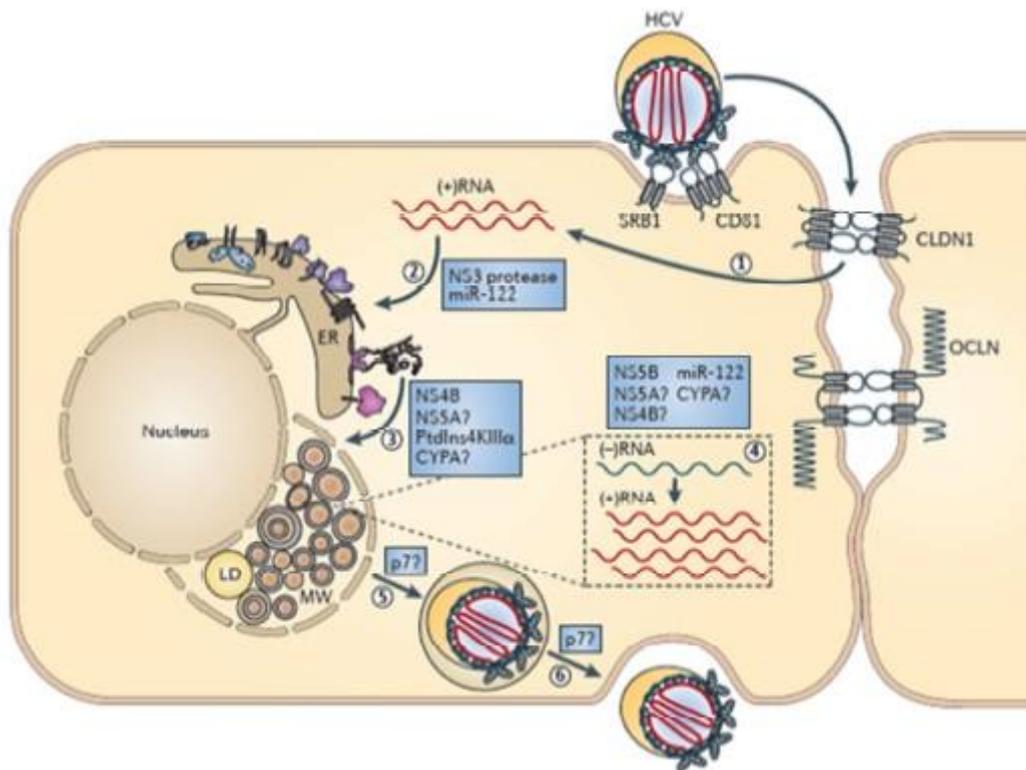


Figura 8. Ciclo replicativo del HCV. Se indican los pasos seguidos durante la replicación del HCV (tomado de Bartenschlager *et al.* 2013).

1.4. Variabilidad Genética del HCV

El HCV tiene una variabilidad de secuencia nucleotídica cuando comparamos todo el genoma viral de aproximadamente 30% entre sus principales genotipos y del 20% en los diferentes subtipos de cada genotipo.

La región 5' NCR es la que contiene el grado de conservación más alto dentro del genoma con un 90% de similitud entre los diferentes genotipos, sin embargo, la región se compone de dominios altamente conservados alternados con dominios variables.

La región del core está relativamente bien conservada, con similitud de secuencias de nucleótidos que oscila entre el 81% y el 88% entre los diferentes genotipos.

La región aminoterminal de E2 constituye una región hipervariable. Es un segmento de 28 aminoácidos que presenta una variación de más del 50% entre las distintas estirpes. Otras regiones muy variables son la región hipervariable 2 de E2 al igual que la NS5A, E1 y E2.

Mientras que las regiones más conservadas son 5'NCR, NS5B y NS3. Se ha estimado una tasa de mutación de aproximadamente $1,4 \times 10^{-3}$ sustituciones de bases por nucleótido por año para todo el genoma (Cuthbert, 1994; Simmonds 2004). La tasa de mutación de la mayoría de los virus de RNA varía de 10^{-1} a 10^{-4} sustituciones por sitio por año, debido a la falta de capacidad de corrección de errores de su polimerasa.

1.4.1. Genotipos

En la última actualización realizada en el año 2014 utilizando genomas completos o casi completos, se resolvieron problemas de nomenclatura para la designación de genotipos y se determinó el uso de un criterio consenso, clasificando al HCV en 7 genotipos y 67 subtipos distintos confirmados (Smith *et al.* 2014). Quedan por resolver aún 22 secuencias de genomas completos las cuales permanecen sin un subtipo asignado.

El Comité Internacional para la Taxonomía de los Virus (ICTV) mantiene y actualiza las estirpes de referencia, sus números de acceso y alineamientos anotados, dicha información se halla disponible en la página Web http://talk.ictvonline.org/ictv_wikis/w/sg_flavi/default.aspx (Figura 9).

Los análisis filogenéticos de secuencias con más del 95% de la región codificante revelaron entonces los 7 principales grupos filogenéticos correspondientes a los genotipos 1-7. Dentro de estos genotipos, el agrupamiento de los diferentes subtipos que los constituyen tiene un soporte de 100% de réplicas de bootstrap (Smith *et al.* 2014).

Basado entonces en el criterio consenso, los subtipos confirmados requieren la secuencia del genoma completo, que tenga diferencia en al menos el 15% de las posiciones nucleotídicas y se cuente con secuencias de al menos otras dos estirpes para la región core/E1 (con más del 90% de la secuencia correspondiente a la posición 869-1292 del genoma de referencia H77, número de acceso AF009606) y la región NS5B (más de 90% de la secuencia correspondiente a la posición 8276-8615 también de H77, número de acceso AF009606) (Smith *et al.* 2014), (Figura 9).

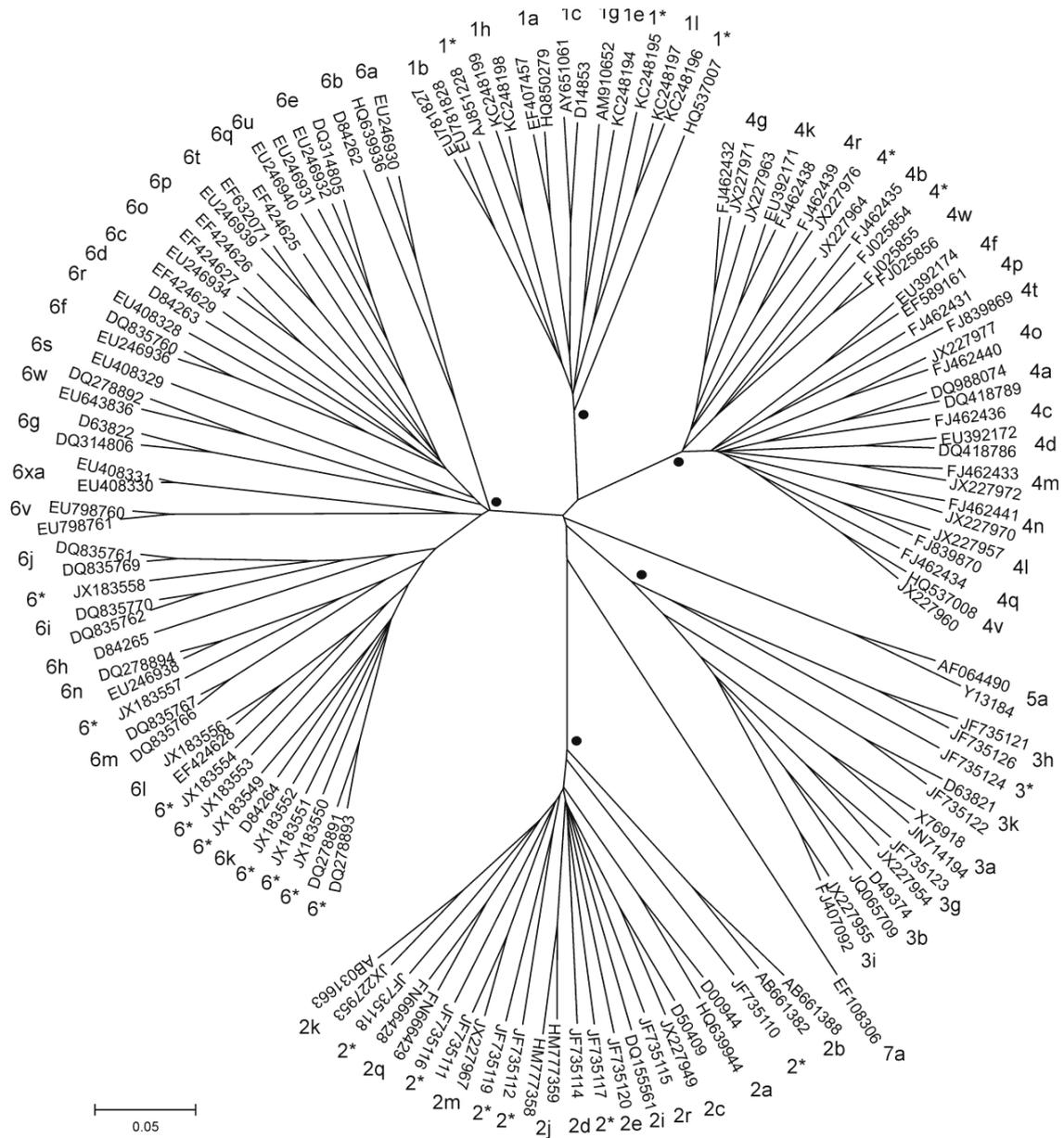


Figura 9. Árbol filogenético de genotipos y subtipos las regiones codificantes de del HCV mediante Neighbor-joinin. Los círculos negros son clados con 100% de bootstrap (Tomado de http://talk.ictvonline.org/ictv_wikis/w/sg_flavi/56.HCV-classification.aspx).

1.4.2. Distribución geográfica de los diferentes genotipos

Si bien HCV es un virus pandémico, los genotipos y subtipos se distribuyen siguiendo un patrón geográfico de circulación. Los genotipos 1 al 3 tienen una distribución global. Los genotipos 4 y 5 se encuentran principalmente en África, y el genotipo 6 se distribuye en Asia (Figura 10).

Las áreas endémicas para los genotipos específicos se encuentran en África Occidental (genotipos 1 y 2), África Centro-Oeste (genotipo 4), el subcontinente indio (genotipo 3), África Central (genotipo 4) y el sudeste de Asia (genotipo 6).

No se ha encontrado un área endémica para el genotipo 5, a excepción de un pueblo en el centro de Francia, donde la infección con HCV 5a de la población local se asoció a vivir en una zona rural llamada Vic-le-Comte (Abergel *et al.* 2007).

Cuando se encuentra una diversidad limitada de los subtipos de HCV en una zona geográfica determinada, ésta puede atribuirse a la reciente introducción de HCV en la población, como se ha documentado para Canadá o Australia. La epidemiología molecular del genotipo 2 apunta a África occidental y central, principalmente a lo largo de la costa atlántica, como su lugar de origen endémico. Se ha encontrado una propagación hacia el este desde la costa de África Occidental a Camerún, que tuvo lugar a lo largo de varios siglos (Markov *et al.* 2009; Lavanchy 2011).

Análisis con reloj molecular datan al ancestro común de HCV en Guinea-Bissau, en torno al año 1470. Aislados de Madagascar y Martinica sugieren que la trata de esclavos histórica y la posible exposición a HCV parenteral durante las campañas de salud pública llevadas a cabo durante la época colonial pueden haber desempeñado un papel en la difusión de los genotipos 2a y 2c (Markov *et al.* 2009).

La distribución geográfica de los genotipos del HCV y la tasa de variación genética son consistentes con la distribución global de HCV, y son compatibles con una larga historia de infección en la mayoría de las poblaciones del mundo (Lavanchy 2011).

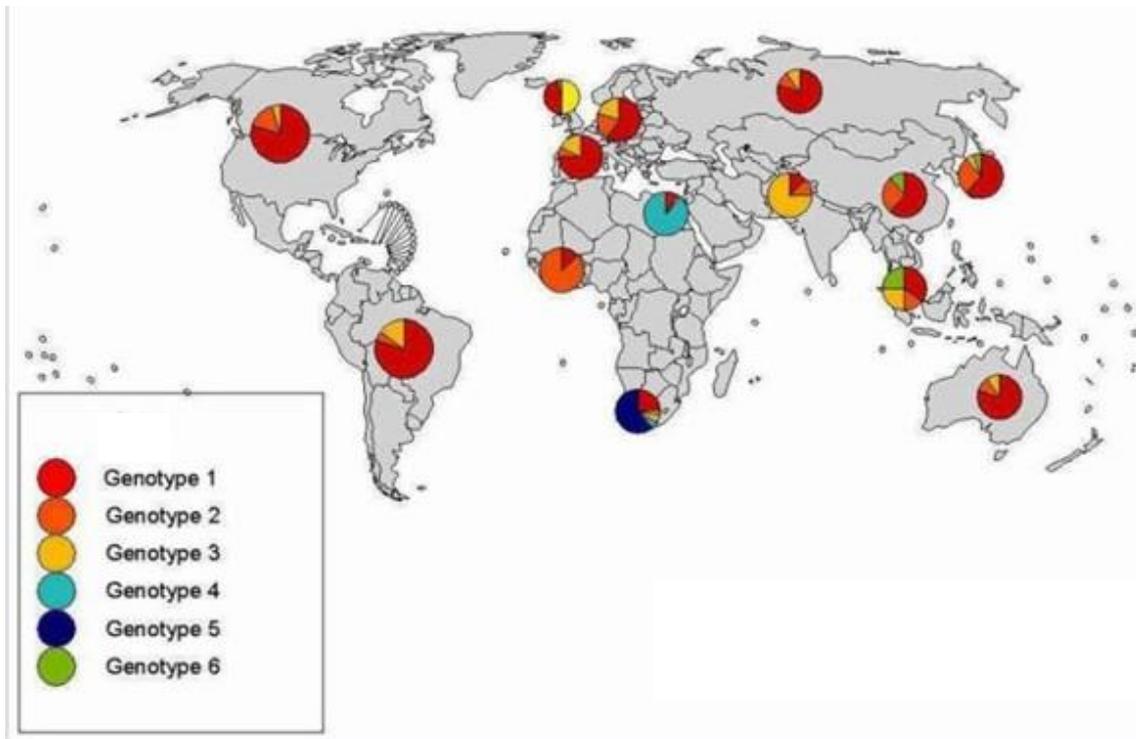


Figura 10. Distribución mundial del HCV. Si bien HCV es un virus pandémico, los genotipos y subtipos se distribuyen siguiendo un patrón geográfico de circulación. Los genotipos 1 al 3 tienen una distribución global. Los genotipos 4 y 5 se encuentran principalmente en África, y el genotipo 6 se distribuye en Asia.

1.5 Evolución del HCV

El tema central en la genética de poblaciones es entender cómo evoluciona una población bajo un conjunto de condiciones dado. La evolución es un proceso continuo en el que una población cambia sus características con el tiempo.

Los cambios genómicos son un prerrequisito para la evolución, los virus han explorado todos los mecanismos a su alcance para obtener variabilidad, los mecanismos moleculares que utilizan son mutación, recombinación, complementación

1.5.1. Mutación

Es un mecanismo clave en la generación de variación genética y nuevos tipos de virus de RNA. Es la alteración localizada y permanente en la secuencia nucleotídica del material genético, es un importante mecanismo para producir nuevas variantes o generar divergencias entre distintas cepas. La mutación puede resultar a partir de la incorporación incorrecta de un nucleótido durante la replicación o de un daño en el ácido nucleico viral.

Las mutaciones pueden ser

a) **Puntuales** cuando una base se cambia por otra, a su vez dependiendo de las bases involucradas se las clasifica en transiciones cuando se cambia una purina por otra purina (A-G) o una pirimidina por pirimidina (C-T); y en transversiones cuando se cambia una purina por una pirimidina o viceversa (A-C; A-T; G-C; G-T); las transiciones ocurren con mayor frecuencia que las transversiones, por su parte las transversiones producen cambios aminoácidos más dramáticos en cuanto a su naturaleza química.

Estas mutaciones pueden producir tres clases de efectos a nivel aminoácido; sustituciones sinónimas (ds) cuando el aa resultante es el mismo, sustituciones no sinónimas (dn) cuando se produce un cambio de aa, y sin sentido cuando se origina un codón stop lo que generalmente da lugar a una proteína inactiva. Esto está asociado a la posición que ocupe el nt en el codón del aa que codifica dado que todas las mutaciones en la 2° posición resultan en sustituciones no sinónimas, mientras que el 96% de las mutaciones en la 1° posición resultan en sustituciones no sinónimas, por su parte las sustituciones en la 3° posición casi siempre resultan en sustituciones sinónimas.

b) Delecciones e inserciones es cuando se pierde o se gana nucleótidos, si esto ocurre en regiones codificantes (en un número distinto a múltiplo de 3) se produce una alteración en el marco de lectura dando lugar a la codificación de aa diferentes.

Una vez originada la variación esta debe de sufrir selección.

Selección: está influenciada por varias clases de fuerzas selectivas:

(a) fuerzas selectivas dependientes del virus, (variaciones que afecten los epitopes de reconocimiento de los receptores celulares o en la replicación del genoma viral).

(b) fuerzas selectivas dependientes del hospedero (respuesta inmune, heterogeneidad fenotípica entre las células de los tejidos donde el virus se replica).

(c) fuerzas selectivas externas (medicamentos antivirales, anticuerpos neutralizantes o vacunación).

Selección positiva y negativa: 2 tipos de selección actúan durante la replicación viral. La relación dn/ds es un indicador del tipo de selección que está actuando en dicha posición, si $dn/ds > 1$ se considera selección positiva si $dn/ds < 1$ la selección es negativa.

La selección negativa (purificadora), actúa eliminando las variantes genómicas con menor capacidad, esta es más intensa en virus bien adaptados al ambiente en que se están replicando, cuando múltiples variantes genómicas infectan una misma célula algunos de estos mutantes deletéreos podrían ser mantenidos por complementariedad.

La selección positiva (direccional) resulta de la dominancia de los genomas que presentan una ventaja evolutiva, como por ejemplo evadir la respuesta antiviral, escapar de los anticuerpos.

En la base de datos de Datamokey existen varios programas que nos permiten predecir si una determinada posición del genoma está siendo sometida a selección positiva o negativa.

Los mutantes de escape son capaces de evadir a los anticuerpos neutralizantes o de reconocer nuevos receptores celulares debido a alguna alteración en los epitopes de reconocimiento de los receptores celulares, produciendo un cambio en el tropismo celular o en el rango de huésped, o de adaptarse a la terapia antiviral.

En el caso de los arbovirus que están expuestos a una alternancia de presiones selectivas dadas por sus 2 hospederos, las que pueden limitar sus tasas evolutivas a pesar de su alta tasa de mutabilidad, la presión selectiva impuesta por el hospedero es más importante pues es allí donde pasan más tiempo.

1.5.2 Recombinación

Es el proceso por el cual 2 genomas parentales intercambian información durante la replicación. La recombinación del HCV sigue siendo un evento raro y el número de casos bien documentados son muy bajos, es probable que los niveles de recombinación en el HCV estén subestimados. La recombinación puede eventualmente facilitar la combinación de mutaciones de resistencia individuales en el mismo genoma, y la vigilancia epidemiológica activa de las formas recombinantes es de interés para la salud pública.

La primera cepa recombinante del HCV se obtuvo en San Petersburgo (Rusia) en 2002 (Kalinina *et al.* 2002). Se trata de un recombinante intergenotipo entre un subtipo 2k y un subtipo 1b. El punto de recombinación fue mapeado en el gen NS2, alrededor de la posición 3175. Posteriormente, el punto de recombinación fue confirmado (Kalinina *et al.* 2004). Esto llevó a la propuesta de una nueva designación de la cepa como RF1 2k/1b, en analogía con la nomenclatura utilizada para formas circulantes. Esta misma cepa recombinante se ha aislado en otros países, Irlanda (Moreau *et al.* 2006); Uzbekistán (Kurbanov *et al.* 2008); Chipre (Demetriou *et al.* 2011); y en Georgia y Francia (Morel *et al.* 2010).

Se han encontrado recombinantes para todos los genotipos del HCV excepto para el genotipo 4. Curiosamente, todos los recombinantes excepto uno el RF_3a/1b, están formados por un extremo 5' de genotipo 2 y extremo 3' de un genotipo diferente, en el que sólo subtipo 1b aparece en más de una de un recombinante (Lee *et al.* 2010). Hasta ahora, sólo cinco casos de recombinación intersubtipos se han descrito, dos implicando subtipos del genotipo 1 y uno entre los subtipos del genotipo 4.

Dos informes de recombinación intrapaciente en el HCV se han publicado hasta la fecha, ambos implican pacientes sometidos a terapia (Moreno *et al.* 2006; Sentandreu *et al.* 2008).

En el genoma existen zonas con diferente susceptibilidad a sufrir eventos de recombinación (hot spots).

Los eventos de recombinación están influenciados por

1) la procesividad de la polimerasa. 2) la existencia de hot spots. 3) la homología de secuencias entre las hebras. 4) la presencia de estructuras secundarias en la zona de recombinación.

La recombinación puede ser vista desde 2 puntos de vista diferente: i) por una parte actúa originando nuevas cepas con mejor fitness dado que conlleva a la formación de genomas mosaicos los que pueden estar mejor adaptados, eso se determina mediante experimentos de crecimiento competitivo. ii) Es también una forma de recuperar regiones que por acumulación de mutaciones se han vuelto muy diferentes Teoría de la “Reina Roja”.

1.5.3 Complementación

La complementación es un proceso por el cual un genoma que expresa una proteína funcional puede promover la replicación de otro genoma estrechamente relacionado cuya proteína correspondiente es defectuosa (o subóptima).

La ocurrencia de complementación entre los componentes de un espectro de mutantes se sugirió en los primeros estudios que mostraron que los clones biológicos individuales aislados a partir de una población de virus poseían valores de aptitud promedio más bajos que toda la población de la que habían sido aislados (Duarte *et al.* 1994).

La complementación entre mutantes virales ha sido ampliamente descrita, y subyace el mantenimiento de los genomas defectuosos (partículas defectuosas) en las poblaciones virales (Ojosnegros *et al.* 2011; Roux *et al.* 1991; Wimmer y Paul 2010).

En la dinámica de las cuasiespecies la complementación puede ser ejercida no sólo entre los genomas dominantes sino también entre los componentes minoritarios de una cuasiespecie, cuyo resultado es la mejora de la aptitud.

1.6 Las cuasiespecies

1.6.1. Teoría de las cuasiespecies

Los virólogos adoptaron el término "cuasiespecies" de una teoría sobre la adaptabilidad de las entidades auto-replicativas, componentes claves en el origen de las formas primitivas de vida. Un estudio pionero de Manfred Eigen es un tratamiento cuantitativo de la evolución de las macromoléculas biológicas (Eigen 1971). El estudio representa el primer tratamiento teórico de la conducta auto-instructiva requerida para la actividad de plantilla, como una necesidad para el origen de la información hereditaria.

Según la teoría de Eigen, una copia maestra de la molécula auto-replicativa produce versiones mutantes con una cierta distribución de probabilidad.

La producción de copias mutantes depende de un factor de calidad que determina la fracción de los procesos de copiado que conduce a una copia exacta de la plantilla y la fracción que conduce a errores en las copias de la plantilla. La teoría fue desarrollada posteriormente por Manfred Eigen y Peter Schuster en una serie de trabajos teóricos, que definían “cuasiespecies” como distribuciones de mutantes de estado estable, dominado por una secuencia maestra poseedoras de la tasa de replicación más alta entre los componentes del espectro de mutantes. La teoría define "hiperciclos" como un principio de la auto-organización natural, al integrar diferentes cuasiespecies en organizaciones de orden superior lo que facilita la evolución hacia formas más complejas (Eigen y Schuster 1978).

Siendo un sistema replicativo con alta tasa mutacional, la teoría de las cuasiespecies establece una condición necesaria para asegurar la conservación estable de la información genética. Tal condición se formuló como un umbral de error. El umbral de error significa que para cualquier complejidad dada (cantidad de información genética no redundante transmitido por un sistema de replicación), existe una tasa de error máxima compatible con el mantenimiento de la información genética. Dicho valor umbral de error depende de la precisión de la replicación y la eficiencia de la secuencia dominante o secuencia maestra respecto al valor medio de la eficiencia de las secuencias restantes de la nube, (Comas *et al.* 2005).

Las dos ecuaciones fundamentales de la teoría de cuasiespecies son: la dinámica de la producción de copias de error y la relación umbral de error.

$$\frac{dx_i}{dt} = (A_i Q_i - D_i)x_i + \sum_{k=1, k \neq i}^n W_{ik} x_k - \Phi_i$$

$$v < v_{\max} = \frac{\ln \sigma_0}{1 - \bar{q}} = \frac{\ln \sigma_0}{\bar{p}} \quad \text{and} \quad \bar{p} < \bar{p}_{\max} = \frac{\ln \sigma_0}{v}$$

La primera ecuación describe la concentración de mutantes i en función del tiempo x_i , y de mutantes k en función del tiempo x_k . Describe la dinámica de la generación de mutantes dentro de la nube de mutantes.

La segunda ecuación describe la relación de umbral de error, donde v_{\max} es la máxima complejidad genética que puede mantenerse durante la replicación y σ_0 es la selectividad o superioridad de la secuencia principal en relación con las restantes secuencias del espectro de mutantes.

Estas ecuaciones están estrechamente relacionadas con las ecuaciones que describen otros modelos de dinámica evolutiva. La teoría de las cuasiespecies formulada inicialmente era determinista, dado que la teoría asume distribuciones de mutantes de tamaño infinito en equilibrio.

Las cuasiespecies se perciben de manera diferente por físicos, químicos, biólogos y médicos expertos en enfermedades infecciosas.

Para los físicos, una cuasiespecies es considerada como una nube en el espacio de secuencias.

Para los químicos, las cuasiespecies son distribuciones de las secuencias de nucleótidos o aminoácidos relacionados pero no idénticos, una definición familiar a la de los virólogos.

Para los biólogos una cuasiespecie es el objetivo de la selección, y el término no implica una modificación al concepto de especie biológica.

Para los expertos en enfermedades infecciosas, las cuasiespecies son los enjambres dinámicos de virus mutantes (particularmente los mutantes resistentes a los medicamentos) que tienen que hacer frente a la terapia antiviral.

Los virólogos utilizan las cuasiespecies virales en el sentido de las distribuciones dinámicas de genomas no idénticos pero estrechamente relacionados, sometidos a un continuo proceso de competencia y selección, y que actúan como una unidad de selección (Domingo 2006; Perales *et al.* 2010).

Estrictamente hablando, las cuasiespecies virales deberían considerarse como una sola unidad de replicación en una célula infectada (Del Portillo *et al.* 2011). Sin embargo, la progenie viral heterogénea a partir de una sola célula invadirá las células vecinas ya sea en un cultivo o dentro de un mismo tejido u órgano in vivo, esto crea un segundo nivel de competencia. En la viremia, se establece la competencia para la invasión de tejidos y órganos entre los virus que se han originado en diferentes unidades replicativas, ya sea en el nivel intracelular o extracelular (Lauring *et al.* 2010; Nowak 2006).

Si bien la teoría de cuasiespecies hace énfasis en el proceso de mutación como la fuente de la variabilidad genómica, se ha ampliado para incluir mecanismos adicionales de variación, tales como recombinación, complementación, duplicación de genes, reordenamiento de segmento de genoma, y de genes (Muñoz *et al.* 2008; Jacobi *et al.* 2006).

1.6.2. Transmisión y divergencia de las cuasiespecies

En los virus que circulan como cuasiespecies, generalmente hay múltiples variantes presentes al momento de la transmisión de la infección. La transmisión de todas las variantes dentro de una cuasiespecie no es uniforme (Gao *et al.* 2002; Weiner *et al.* 1993). A menudo, la transmisión resulta en un cuello de botella para la población, donde sólo una pequeña fracción de las variantes presentes en la nube original de la cuasiespecie pasa al nuevo hospedero. Las variantes dominantes en el inóculo a menudo se adaptan mal al nuevo entorno. Como resultado, una variante de menor importancia en la cuasiespecie original, a menudo se convierte en dominante en el nuevo hospedero.

Cuando múltiples variantes hacen con éxito la transición al nuevo hospedero, hay una gran diversidad de cuasiespecies en la fase temprana de la infección (Herring *et al.* 2005).

Un estudio de un brote en el que los pacientes estaban infectados de una sola fuente mostró que las variantes presentes en el inóculo pueden dar lugar a poblaciones divergentes dentro de cada paciente. Las dos poblaciones pueden ser tan diferentes entre los pacientes como lo son con las secuencias de la progenie de otros pacientes (Ray *et al.* 2005).

1.6.3 Compartimentación de las cuasiespecies

Los Análisis de variantes aisladas a partir de diferentes compartimentos del cuerpo muestran que los miembros de las cuasiespecies no están distribuidos al azar.

Variantes con diferentes tropismos celulares y compartimentación de los genomas se han observado para un número de virus de RNA, incluyendo VIH y HCV.

Se han encontrado variantes de secuencia que están restringidos a un compartimento particular del cuerpo, en el suero (Cabot *et al.* 2000) en el Sistema Nervioso Central (Fishman *et al.* 2008; Forton *et al.* 2004).

Se ha descrito un paciente infectado con genotipos 1a y 1b del HCV. Variantes genotipo 1b se encuentran exclusivamente en el hígado, mientras que las variantes de genotipo 1a eran encontrado en el hígado, plasma y tejido cerebral (Fishman *et al.* 2008).

La demostración de la compartimentación de las secuencia de una sola cuasiespecies se realiza por lo general con la construcción de un árbol filogenético,

donde las variantes de un determinado compartimiento se agrupan juntas, a pesar de que la segregación completa de variantes por compartimentos en clados es rara.

Se ha demostrado que las distancias genéticas observadas entre las variantes aisladas de diferentes compartimentos son significativamente mayores que las distancias genéticas observadas entre las variantes en el mismo compartimiento (Roque Afonso *et al.* 1999; Poss *et al.* 1998).

1.6.4. Métodos de análisis de cuasiespecies

Para el análisis de una cuasiespecie es útil comenzar estableciendo la secuencia consenso de todas las variantes de la población. El RNA viral aislado a partir de una muestra clínica o a partir de un sistema experimental, es amplificado por RT-PCR, y el amplicón se secuencia directamente, el análisis del cromatograma nos permitirá obtener la secuencia consenso.

Para la obtención de las secuencias de las variantes individuales, es conveniente que el producto de PCR se clone en un plásmido, el que se puede utilizar para transformar bacterias. Las bacterias transformadas se colocan en placas para permitir la selección de colonias individuales, que se supone que se que presenten un único plásmido, por tanto, sólo una variante viral.

Mediante la recuperación y secuenciación del ADN plasmídico de una colonia bacteriana individual, se puede obtener la secuencia de una sola variante. Las opiniones varían en el número de clones que deben ser secuenciados para investigar adecuadamente las cuasiespecies.

La comparación de la secuencia consenso de una población con las secuencias de los componentes individuales de la misma población indica la dinámica de cuasiespecies, la que se refleja básicamente en las variaciones de frecuencia (capacidad relativa) de subconjuntos de genomas en respuesta a los cambios ambientales.

El reciente desarrollo de la secuenciación masiva ha permitido a los investigadores secuenciar una fracción mucho mayor de las cuasiespecies. Este aumento en los genomas de la muestra se produce a expensas de una menor cobertura en la longitud de las secuencia (menos de 200 pares de bases), (Eriksson *et al.*, 2008).

Una vez que se obtienen las secuencias de las variantes, varios tipos de análisis bioinformáticos se puede realizar.

Se puede estudiar la proporción de sustituciones sinónimas y no sinónimas, mutaciones sinónimas o silenciosas son aquellas que no cambian la secuencia de aminoácidos en la proteína, son consideradas tradicionalmente como evolutivamente neutral. Sustituciones no-sinónimas son aquellas en las que cambia la secuencia de aminoácidos de la proteína. La relación de sustituciones no sinónimas por sitio elegibles (dn) y sustituciones sinónimas por sitios elegibles (ds) es un indicador de la fuerza de la selección positiva o negativa que actúa sobre la población.

Un cociente dn/ds mayor que 1 es un indicativo que la "selección positiva" actúa sobre la población (Kimura, 1977). La selección positiva es el proceso por el cual un genotipo (o conjunto de genotipos) se convierte en dominante durante la evolución de la población. En las cuasiespecies virales, una distribución de mutante cuyos componentes comparten rasgos seleccionables, esa distribución de mutantes se convierte en dominante, (Perales *et al.* 2005).

Por el contrario, una proporción de menos de 1 indica "selección negativa". La selección negativa es el proceso por el cual un genotipo (o conjunto de genotipos) se elimina durante la evolución de la población.

Cabe destacar que la distinción entre selección positiva y negativa puede llegar a ser confusa porque en un proceso de competencia, la selección negativa de subconjuntos de genomas puede tener el mismo resultado que la selección positiva de los subconjuntos restantes y viceversa, además, en la selección negativa no tiene por que resultar la eliminación de los subconjuntos de genomas, ya que pueden dar lugar a su mantenimiento en bajas frecuencias.

Estos argumentos conducen a la necesidad de cuantificar ventajas selectivas relativas de las poblaciones virales a través de las mediciones de adaptabilidad (fitness).

La adaptabilidad es un parámetro importante en la genética evolutiva, que también se incluyó en la formulación de la teoría de cuasiespecies y en la ecuación umbral de error. Fue adaptado por los virólogos para cuantificar la capacidad de replicación relativa de un virus, generalmente medido en experimentos de competición con el crecimiento de un virus aislado de referencia, ya sea en cultivo celular o in vivo (Quiñones-Mateu y Arts 2006). Mediciones de aptitud han sido comparadas a la determinación de un coeficiente de selección (Maree *et al.* 2000).

Comparación de la secuencia de nucleótidos consenso de una población con los de los componentes individuales de la misma población indica que la dinámica de

cuasiespecies se refleja básicamente en las variaciones de frecuencia (aptitud relativa) de subconjuntos de genomas en respuesta a los cambios ambientales.

Cuando un espectro de mutantes alberga un amplio repositorio de variantes, la selección de subconjuntos del genoma en un nuevo entorno puede ser muy rápida.

Tamaño de la población, la heterogeneidad genética, la capacidad de adaptación, y la tasa de evolución dependen de un conjunto interconectado de parámetros, algunos susceptibles de cuantificación y otros difíciles de medir. Parámetros importantes son la velocidad de la multiplicación del genoma, la tasa de mutación, y la tolerancia de los genomas a aceptar mutaciones y permanecer funcional. La tasa de mutación, la tolerancia a las mutaciones (mutaciones no letales con un rango de valores de aptitud), y el tamaño de la población determinarán la amplitud del espectro de mutantes y el paisaje de mutaciones minoritarias y su frecuencia.

Otros parámetros de interés común en el estudio de la composición de cuasiespecies son la complejidad y la diversidad de la población.

Complejidad se refiere al número de diferentes secuencias presentes en la población. Por lo general, la complejidad se presenta como la proporción de variantes únicas entre el número total de clones analizados como la entropía de Shannon:

$$-\frac{\sum_{i=1}^n f_i (\ln f_i)}{N}$$

Donde n es el número diferentes de especies identificadas, f_i es la frecuencia observada de una variante en particular en la cuasiespecie, y N es el número total de clones secuenciados. Se divide entre N para normalizar, por las diferencias en el número de clones analizados.

Un amplio espectro de variantes (alta complejidad) indica una capacidad similar entre las variantes virales y una estabilidad de la población. Del mismo modo, la baja complejidad puede reflejar diferencias más grandes en la aptitud y un cambio más reciente en el medio. Alternativamente, baja complejidad puede reflejar una falta de presión selectiva.

La diversidad de una cuasiespecie se refiere a la relación de los individuos dentro de la población. La medida más básica de la diversidad es la distancia media de Hamming, es decir, el número de posiciones mutadas en una secuencia particular con respecto a una secuencia dominante, secuencia de consenso, u otra referencia.

La diversidad se mide usando un algoritmo de métrica de distancia genética tales como Jukes-Cantor o Kimura-2-parámetro, que tienen en cuenta factores adicionales, tales como las probabilidades de transición (de purina a purina o pirimidina a pirimidina), de transversiones (pirimidina a purina y viceversa), y la composición global de aminoácidos o de nucleótidos en las secuencias. Las distancias genéticas dentro y entre las cuasiespecies se utilizan a menudo para generar múltiples secuencia de alineaciones y los árboles filogenéticos. Las secuencias que se agrupan o se asignan a los mismos subtipos están más estrechamente relacionadas entre sí.

La fuerza de la agrupación de secuencias en un árbol filogenético se puede evaluar mediante la aleatorización repetidamente las secuencias y volver a crear el árbol (bootstrapping).

1.6.5. Genomas memoria

El aumento de la adaptabilidad resultante de la replicación durante la selección permite el posterior mantenimiento de la subpoblación seleccionada como genomas memoria (Ruiz-Jarabo 2000). Estos pueden permanecer presente en el espectro de mutantes a frecuencias más altas que las dictadas por las tasas de mutación basales (Briones y Domingo 2008; Ruiz-Jarabo 2000).

La búsqueda de la memoria en las cuasiespecies se inspiró en el hecho de que la movilización de elementos biológicos minoritarios en respuesta a un estímulo es típico de los sistemas adaptativos complejos, tales como el sistema inmunológico.

Se ha postulado que los genomas presentes en el espectro de mutantes de cuasiespecies virales pueden incluir genomas que representan un registro de los genomas que eran dominantes en las fases anteriores del mismo linaje evolutivo.

Los genomas memoria le brindan a la población viral la capacidad para responder a las restricciones selectivas ya experimentados por el mismo linaje evolutivo (Domingo 2000; Perales *et al.* 2010). La dinámica de la adquisición de la memoria se espera que operen in vivo cuando un agente antiviral se utiliza para tratar una infección y se selecciona un mutante resistente a los antivirales que luego puede ser eliminado por otros componentes del espectro mutante. La conversión en genoma memoria puede ocurrir como resultado de la interrupción del tratamiento o aplicación de un régimen de tratamiento alternativo.

1.7. Representación de historias evolutivas

La historia evolutiva de un conjunto de taxones se suele representar por un árbol filogenético, y este modelo ha facilitado la discusión y pruebas de hipótesis. Sin embargo, es bien sabido que escenarios evolutivos más complejos son mal descritos por tales modelos. Aún cuando la evolución se suceda a manera de árbol, los datos pueden ser también visualizados mediante el uso otros métodos, tales como las redes filogenéticas.

La reconstrucción de filogenias de cuasiespecies suele ser una ardua tarea, dado los importantes tamaños de las muestras y a las pequeñas distancias génicas entre los genomas; la multitud de posibles árboles se puede expresar mejor por una red que muestre los posibles caminos evolutivos en forma de ciclos.

1.7.1. Árboles filogenéticos vs redes de haplotipos

Un árbol filogenético se define comúnmente como un grafo acíclico con forma de árbol, con hojas etiquetadas que representa la historia de la evolución de un conjunto de taxones, generalmente con longitudes de rama, ya sea con raíz o no.

Los árboles filogenéticos, basados en nucleótidos o aminoácidos, son grafos acíclicos que se utilizan para representar relaciones filogenéticas entre especies (Huson *et al.* 2011). Cada nodo tiene 1 grado de entrada y al menos 2 grados de salida.

El concepto de una red filogenética no está tan bien definido, y existen muchos usos diferentes del término. Una fuente importante de confusión ha sido que diferentes autores definen la red filogenética como algún tipo particular de red analizada en su estudio. Por ejemplo, un artículo sobre recombinación (Gusfield y Bansal 2005) define una red filogenética como una red de recombinación, mientras que un artículo referente a hibridación, (Linder y Rieseberg 2004) definen una red filogenética como una red de hibridación.

Se puede definir una red filogenética como cualquier red en el que están representados los taxones por nodos y sus relaciones evolutivas están representadas por los bordes. Las redes filogenéticas de haplotipos (o redes alélicas), son grafos con la presencia de ciclos, donde los nodos representan diferentes haplotipos y la longitud de las ramas representa la diferencia en números de nucleótidos entre ellos. Estas redes pueden tener hasta 2 grados de entrada para cada nodo.

Las redes filogenéticas son una generalización de los árboles filogenéticos que se puede utilizar para mostrar historias evolutivas más complejas, incluyendo eventos reticuladas, tales como hibridaciones, recombinaciones y la transferencia horizontal de genes (Leo van Iersel *et al.* 2010).

1.7.2. Tipos de Redes

Se puede distinguir diferentes tipos de redes:

Un primer tipo lo constituyen los árboles filogenéticos (Figura 11a y 11b).

Un segundo tipo, es la red de división, que es obtenida como una combinatoria de árboles filogenético y está diseñada para representar incompatibilidades dentro y entre los conjuntos de datos (Figura 12a).

Un tercer tipo, son las redes reticuladas, que representa historias evolutivas en presencia de eventos reticulados tales como la hibridación, la transferencia horizontal de genes, o los eventos de recombinación (Figura 12b). Existen otros tipos de redes (Posada y Crandall 2001), para representar la duplicación de genes (Hallett y Lagergren 2000; Durand *et al.* 2005), la coevolución hospedero-parásito (Charleston 1998).

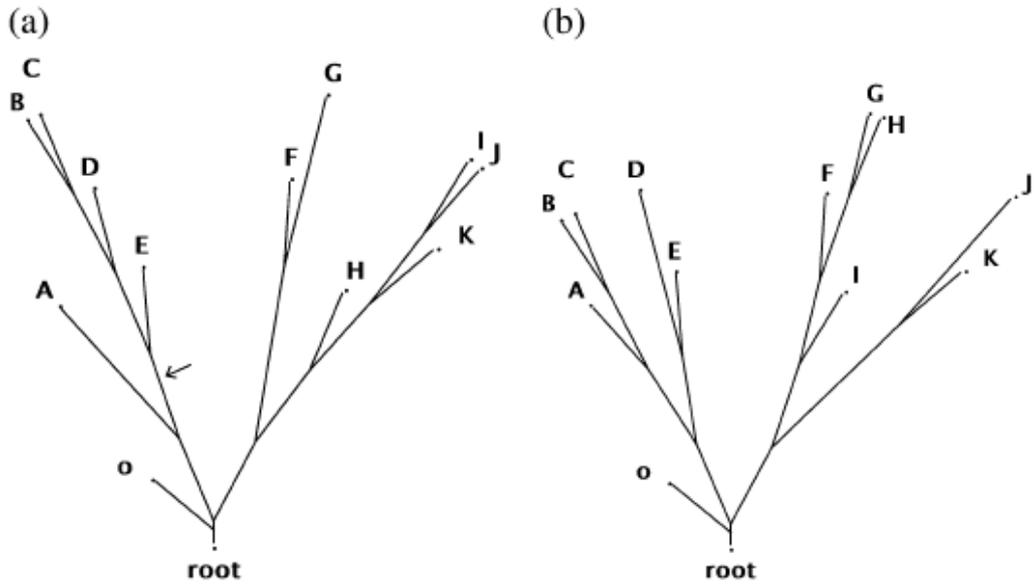


Figura 11. Representa 2 posibles árboles filogenéticos de un mismo análisis. La flecha señala el punto en donde se divide a los taxones en dos posibles agrupaciones (B, C, D, E) y (o, A, F, G, H, I, J, K).

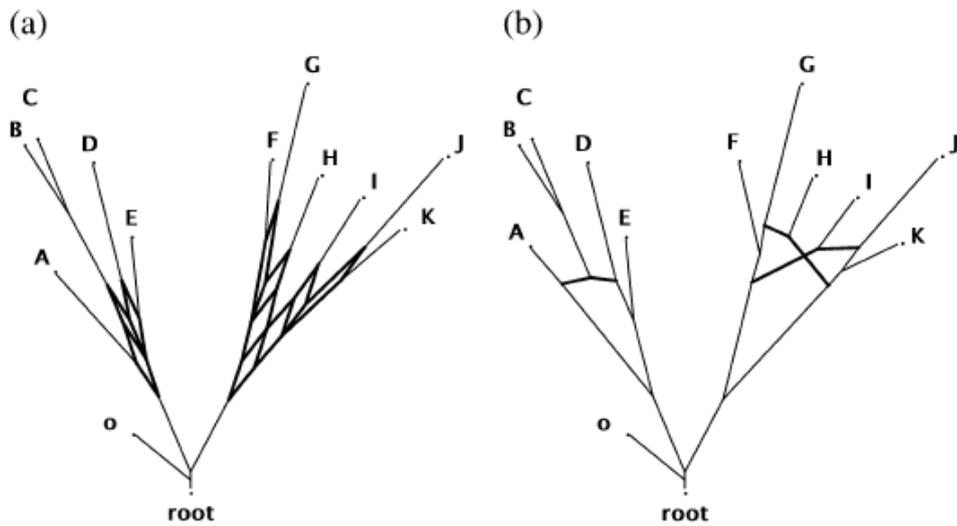


Figura 12. Dos tipos diferentes de redes filogenéticas.
 (a) Una red de división que representa todas las divisiones presentes en los dos árboles representados en la anterior figura. Aquí, cada banda de bordes paralelos corresponde a una rama contenida en uno de los árboles de entrada. Los nodos no corresponden necesariamente a antepasados hipotéticos.
 (b) Una red reticulada que explica los dos árboles postulando tres reticulaciones que dan lugar a los clados (B,C) (H), e (I). Esta red describe explícitamente una historia evolutiva probable, los nodos internos corresponden a taxones ancestrales, y los bordes representan los descendientes.

1.7.2.1. Árboles Filogenéticos

Un árbol filogenético corresponde a una colección de divisiones compatibles con los pesos y longitudes.

El análisis de árbol filogenético tiene por objeto encontrar el árbol filogenético que mejor explica los patrones observados en los datos. El método de generación de árboles intentará colocar un árbol, incluso si existe un gran desfase entre los datos y el mejor árbol que el método puede encontrar. Esta brecha es la principal causa de artefactos de reconstrucción (Steel 2005).

1.7.2.2. Las redes de división

Las redes de división se utilizan para representar señales incompatibles y/o ambiguas en un conjunto de datos. En la red, se presentan bordes en paralelo, en lugar de las ramas individuales, se utilizan para representar las escisiones calculadas a partir de los datos. Cada borde se asocia con una división de los taxones, pero puede haber un número de bordes paralelos asociados con cada división. Por lo tanto, estas redes de división sólo proporcionan una representación implícita de la historia evolutiva, (Morrison 2005).

La distancia entre dos taxones en una red de división se define como la suma de los pesos o longitudes de los bordes a lo largo del camino más corto entre los taxones (Bryant y Moulton 2004). Esta distancia se puede calcular directamente de los pesos asociados a las divisiones y no cambia para diferentes representaciones de redes abiertas.

Cada red división representa una colección única de divisiones. Sin embargo, la singularidad no se sostiene en la otra dirección, un conjunto determinado de divisiones puede tener varias diferentes representaciones de la red. Debido a esta no unicidad, es inapropiado considerar los nodos internos como ancestros hipotéticos (Bandelt 1992).

1.7.2.3. Las redes reticuladas

Mientras que las redes de división proporcionan una imagen implícita de la historia evolutiva, se emplean conjuntos de bordes paralelos para representar divisiones, nodos anteriores o internos en una red de división no se corresponden necesariamente

con antepasados hipotéticos, las redes reticuladas proporcionan una visión explícita de la evolución, en este tipo de red, los bordes representan linajes de descendencia o eventos reticulados tales como la hibridación, la transferencia horizontal de genes, o la recombinación, y todos los nodos corresponden a los ancestros hipotéticos, estas redes reticuladas suelen ser enraizadas, por lo que los bordes tienen una dirección con un significado evolutivo.

Este tipo de redes proporcionan una representación explícita de la historia evolutiva, representada generalmente como un árbol filogenético con bordes adicionales, donde los nodos internos representa la especie ancestral y nodos con más de dos padres corresponden a eventos reticulares tales como hibridación o recombinación.

2. Materiales y métodos

2.1. Construcción de los Datasets.

Para la realización de nuestros objetivos se construyeron varios datasets, a partir de secuencias obtenidas de la base de datos de National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov>; y de HCV Database <http://HCV.lanl.gov>.

El dataset 1.- Se construyó en base a secuencias de la región NS5A del genotipo 1b de HCV tomadas de los trabajos de Puig-Basagoiti. Correspondientes a 4 pacientes sometidos a terapia con INF α /RVB, administra por vía subcutánea, durante 24 semanas, con diferentes respuestas al tratamiento y un paciente crónico no tratado. Los pacientes fueron estudiados en la primera, segunda y cuarta semana de tratamiento, (Puig-Basagoiti *et al.* 2005).

Dataset 2.- Las secuencias utilizadas corresponden a la región NS5A del genotipo 3a de HCV, tomadas de los trabajos de Bittar (Bittar *et al.* 2010; Bittar *et al.* 2013). Pertenecen a pacientes sometidos a terapia con INF α /RVB, los que se clasificaron en: pacientes que no respondieron a la terapia, pacientes que presentaron respuesta viral al final del tratamiento, y pacientes que presentaron respuesta viral sostenida. Las secuencias abarcan desde antes de la terapia hasta 5 meses después de finalizada la misma. También integran este dataset secuencias tomadas de los trabajos de Jardim corresponden a la región NS5A del genotipo 1a de HCV, (Jardim *et al.* 2013).

Dataset 3.- Se construyo en base a secuencias tomadas de los trabajos de Farci. Correspondientes a la región del gen de la envoltura E1/E2 de HCV que incluye la región hipervariable 1. El paciente adquirió la infección por transmisión materno-fetal, no siendo tratado, se estudió desde su nacimiento y durante 90 meses (Farci *et al.*, 2006).

Dataset 4.- Utilizando secuencias de HCV de las regiones NS5A y E1/E2 del genotipo 1a, infectando pacientes no tratados estudiados desde la fase aguda a la cronicidad (16-18 años) tomadas de los trabajos de Ramachandran, (Ramachandran *et al.* 2011).

Dataset 5.- En base a secuencias de cuasiespecies intra-hospedero, correspondientes a la región HVR1 que circula en un paciente crónico infectado con HCV de genotipo 1b, obtenido de los trabajos de Kamila Carballo-Cortés, por ultrasecuenciación profunda con ROCHE 454 (Kamila Carballo-Cortés *et al.* 2013).

2.2 Alineación de las secuencias

Para la realización de los análisis las secuencias fueron alineadas mediante MEGA 5.05 (Tamura *et al.* 2011), se utilizó el servidor web Datamonkey (Delpont *et al.* 2010), para identificar el modelo evolutivo que mejor se ajuste a nuestros datos.

2.3. Análisis de uso de codones

Para el análisis de uso de codones comparativo entre los pacientes se utilizó el Programa CODON W, que permitió determinar el Uso Relativo de Codones Sinónimos (RSCU), (Wright 1990).

2.4. Análisis coalescente bayesiano de Cadenas Monte Carlo de Markov (MCMC)

Este análisis constituye una medida de la evolución de una población de virus de RNA, lo que permite la reconstrucción de la dinámica de poblaciones en el pasado sobre la base de datos de secuencias moleculares. Esta reconstrucción se realiza con frecuencia en un marco bayesiano para dar cuenta de la incertidumbre en el árbol filogenético subyacente (Drummond *et al.* 2005). El análisis coalescente bayesiano de Cadenas Monte Carlo de Markov (MCMC) es actualmente, un análisis estándar utilizado para la reconstrucción de la dinámica evolutiva de las poblaciones ancestrales (Drummond y Rambaut 2007). Por estas razones hemos utilizado un enfoque bayesiano MCMC para estudiar el modo de evolución de la población de cuasiespecies del HCV tal como se aplica en el paquete BEAST 2 (Stadler *et al.* 2013).

Como modelo de población se utilizó el skyline bayesiano clásico y el de nacimiento-muerte. El modelo de skyline de nacimiento-muerte esencialmente combina dos enfoques anteriores. Un modelo de skyline que asume que las muestras fueron tomadas en un momento en el tiempo (Stadler. 2011), y un modelo de muestreo secuencial para las tasas epidemiológicas constantes (Stadler *et al.* 2010). La combinación de estos dos modelos, en un marco de inferencia bayesiana, da lugar al modelo nacimiento-muerte (Stadler *et al.* 2012). Este nuevo enfoque fue empleado con éxito para el estudio de muestras de HCV que fueron recogidas en un punto de tiempo (Stadler *et al.* 2013).

Se emplearon entre 30 y 90 millones de generaciones de MCMC (según los datasets). La incertidumbre estadística en los datos se refleja en más del 95 % de densidad de probabilidad valores (HPD). Los resultados se analizaron utilizando el programa v1.4 TRAZADOR (Drummond y Rambaut 2007), en el paquete BEAST. La convergencia se evaluó con los valores de SEE (tamaño efectivo de la muestra).

2.5. Análisis filogenético

Para los análisis filogenéticos se emplearon 2 enfoques diferentes.

En primer lugar, los árboles de máxima credibilidad, los que fueron generados usando el programa Tree Annotator del paquete BEAST y el programa FIGTREE v1.2.2 (disponible en: <http://tree.bio.ed.ac.uk>).

En segundo lugar, el análisis de conjuntos de árboles, para lo cual se utilizó el programa DensiTree (Bouckaert 2010). DensiTree es un programa para la elaboración de conjuntos de árboles almacenados en formato Nexus. La idea principal es analizar a todos los árboles en el set, pero en lugar de utilizar líneas opacas, utiliza la transparencia. Como resultado, en las áreas donde muchos de los árboles concuerdan en longitud y topología de la rama, habrá muchas líneas dibujadas y la pantalla mostrará un área densamente coloreada. Las áreas en las que hay un par de topologías en competencia serán resaltadas por una red de líneas. La incertidumbre en las alturas de los nodos y su distribución puede demostrarse mediante frotis de todo el nodo de altura media (Bouckaert 2010). Un total de 20.000 árboles se extrajeron utilizando este enfoque para construir el árbol que se muestra en este trabajo, utilizando la salida del análisis Bayesiano MCMC.

2.6. Análisis de redes

Las redes filogenéticas son una generalización de los árboles filogenéticos que se puede utilizar para mostrar historias evolutivas más complejas, incluyendo eventos reticulados, tales como hibridaciones, recombinaciones y la transferencia horizontal de genes.

La reconstrucción de filogenias de cuasiespecies suele ser una ardua tarea, debido a los importantes tamaños de las muestras y a las pequeñas distancias génicas

entre los genomas; la multitud de posibles árboles se puede expresar mejor por una red que muestre los posibles caminos evolutivos en forma de ciclos.

Para inferir las relaciones intraespecíficas entre haplotipos fue usado el algoritmo de "Red Mediana de Uniones" (MJ) (por su nombre en inglés, Median Joining Network) Brandelt *et al.* 1999). El método, se basa en la combinación de características del algoritmo de Kruskal para encontrar árboles de máxima verosimilitud, (Kruskal 1956) y del algoritmo heurístico de Farris de máxima parsimonia (MP), (Farris 1970) que añade secuencialmente nuevos vértices llamados "vectores medios", y añade la no resolución de las homoplasias que se generen. El método de MJ esta relacionado con los enfoques para la estimación de máxima parsimonia ajustando el nivel de homoplasia, mediante el parámetro " ϵ "; siendo éste método aplicado a caracteres multivariados. Esta aproximación para estimar genealogías intraespecíficas toma en cuenta fenómenos propios del nivel poblacional como la coexistencia en una genealogía, de genes ancestrales y derivado, y la ocurrencia de eventos de recombinación. Dichos eventos producen relaciones reticuladas, que no siempre son bien representadas en árboles que se bifurcan.

Para implementar el método de Median Joining Network utilizamos el software Network 4.5.1.6 (<http://www.fluxus-engineering.com/>), a partir del número de sustituciones pareadas. Además el criterio de optimización, Máxima Parsimonia, introduce vectores medios que representan haplotipos intermediarios ausentes (Polzin *et al.* 2003).

Para ello utilizamos los parámetros fijados en el software por defecto, seleccionando el algoritmo Median Joining (Brandelt *et al.* 1999), por tratarse de datos nucleotídicos, que son multiseriados, se utilizo u valor de $\epsilon = 0$ (ponderación de distancia genética).

Así mismo para reducir la complejidad de la red y evidenciar mejor la relación entre los haplotipos, se selecciono previo al procesamiento de los datos, la opción "Star Contraction" (Forster *et al.* 2001) que además contribuye a evidenciar topologías tipo estrella, característica de eventos de expansión poblacional.

Ambos tipos de representación, filogenia y redes nucleotídicas muestran la misma información, excepto que cuando aparecen homoplasias (conexiones ambiguas), la filogenia la soluciona, teniendo en cuenta a los árboles mas parsimoniosos por separado y luego construyendo un árbol consenso, en el que se derriban clados creando

un árbol multifurcado con pérdida de información; en cambio la red lo que hace es construir redes de haplotipos con ciclos entre sus secuencias, las redes permiten representar todas las conexiones, incluyendo las ambiguas en una sola figura, en forma de ciclos.

La característica principal que distingue a los árboles de las redes es la posibilidad de incluir ciclos, la red permite representar todas las conexiones ambiguas en una sola figura, en conclusión cuando las conexiones alternativas son igualmente parsimoniosas, los gráficos de red de haplotipos parecen ser mas apropiados para representar las variaciones de secuencias nucleotídicas, que los filogramas consensos.

Los prerequisites para el empleo de MJ son:

- a) Las secuencias deben de estar correctamente alineadas.
- b) Los datos ambiguos son infrecuentes y no existe recombinación.
- c) La medida de distancia entre dos secuencias esta dada por el conteo de caracteres diferentes, (distancia de Hamming).

En nuestros estudios encontramos redes que presentan un comportamiento de tipo disassortative, en que los nodos de grado elevado están mayormente conectados a nodos de grado menor que el propio, (Newman 2002).

2.7. Identificación de epítopes

Con el fin de identificar epítopes a lo largo de las secuencias analizadas se utilizó el programa BepiPred, del servidor Datamonkey (Delpont *et al.* 2010). Se mapearon las mutaciones encontradas con la finalidad de asociarlas a los epítopes predichos.

2.8. Mapeo de sustituciones aminoacídicas.

A fin de identificar la ubicación las sustituciones aminoacídicas, se mapearon las sustituciones aminoacídicas encontrada con la estructura correspondiente tomada de la Protein Database, mediante el programa PDB ProteinWorkshop 3.6 (Moreland *et al.*, 2005).

3. Objetivos

3.1 Objetivo general

- A) Analizar la dinámica evolutiva de las cuasiespecies intrahospedero.
- B) Investigar miRNAs asociados al HCV.

3.2 Objetivos específicos

- 1) Analizar la dinámica evolutiva de las cuasiespecies en función de la enfermedad y la respuesta a la terapia.
- 2) Analizar dinámica evolutiva de las cuasiespecies en distintas regiones del genoma en un mismo paciente crónico sin terapia.
- 3) Analizar la dinámica evolutiva de las cuasiespecies durante la transmisión materno fetal.
- 4) Analizar la dinámica evolutiva de las cuasiespecies por pirosecuenciación.

4. Objetivo específico 1.

Analizar la dinámica evolutiva de las cuasiespecies en función de la enfermedad y la respuesta a la terapia.

Dos importantes cuestiones sin resolver en la patogénesis de la enfermedad por el HCV son; por qué los resultados de la infección crónica por HCV varía tan ampliamente de paciente a paciente y cuáles son los mecanismos que subyacen a los diferentes resultados clínicos.

Mientras que un tercio de los pacientes infectados tienen una enfermedad hepática activa y progresiva que conduce a la cirrosis y HCC, otros dos tercios tienen una enfermedad leve, no progresiva.

El hecho de que pacientes con el mismo genotipo y carga viral respondan de forma diferente al tratamiento llevo a buscar factores virales y del hospedero que pueden ser responsables de la eficacia de la terapia.

En cuanto a los factores virales, hay consenso en que los niveles de RNA en suero y el genotipo viral se correlacionan pobremente con el desarrollo de la enfermedad (Alter y Seeff 2000; Yeo *et al.* 2001).

El actual estándar de tratamiento para la hepatitis C crónica es INF pegilado α en combinación con la Rivabirina, (INF α /RBV).

El tratamiento optimizado con estos agentes puede inducir una respuesta virológica sostenida en 70% a 80% de los pacientes con el genotipo 2 y 3, pero en sólo 40% a 50 % de los pacientes infectados con el genotipo 1 (Afdhal *et al.* 2011).

Mientras que algunos informes sugieren una correlación entre el grado de variabilidad de las cuasiespecies del HCV y la gravedad de la enfermedad hepática (Koizumi *et al.* 1995; Yuki *et al.* 1997) otros no lo hicieron (Rothman *et al.* 2005).

En la mayoría de estos estudios se analizó un único punto temporal durante el largo curso de la enfermedad, los estudios longitudinales proporcionan una imagen más fiable de la relación entre la evolución viral y la progresión de la enfermedad

Para analizar la dinámica evolutiva de las cuasiespecies de HCV intrahospedero en pacientes sometidos a terapia en función de la respuesta a la misma construimos dos datasets.

Un dataset (A) para estudiar la dinámica evolutiva de la cuasiespecie durante las primeras 4 semanas de terapia, constituido con secuencias tomadas del trabajo de Puig-Basagoiti de 2005 correspondientes a la región de la NS5A de HCV, (nt 6786 a 7547 en referencia a la cepa H77c; AF011751), de pacientes crónicos infectados con HCV genotipo 1b sometidos a terapia combinada de INF α /Ribavirina durante 24 semanas y pacientes no sometidos a terapia (Puig-Basagoiti *et al.* 2005).

Dicho dataset consta de cinco pacientes; dos pacientes no respondedores (NR), en los que el RNA viral fue detectable por PCR durante todo el tratamiento, (pacientes 10 y 12); dos pacientes con respuesta viral sostenida (RS), en los que el RNA viral dejó de ser detectable por PCR luego de la cuarta semana de tratamiento, (pacientes 7 y 9); y un paciente crónico sin tratamiento (C), (paciente 16).

Otro dataset (B) para estudiar la dinámica evolutiva antes, durante y luego de finalizada la terapia, constituido con secuencias tomadas de los trabajos de Bittar y Jardim correspondientes a la región NS5A de pacientes crónicos infectados con HCV genotipo 3a, sometidos a terapia combinada de INF α /Ribavirina durante 24 semanas, (Bittar *et al.* 2010; Bittar *et al.* 2013); y otros pacientes crónicos infectados con HCV genotipo 1a, también sometidos a terapia combinada de INF α /Ribavirina durante 24 semanas, (Jardim *et al.* 2013).

Dicho dataset consta de 5 pacientes con genotipo 3a, de los cuales tres pacientes son no respondedores (NR), en los que el RNA viral fue detectable por PCR durante

todo el tratamiento, (pacientes: P07, P75 y P145); y 2 pacientes que presentan respuesta viral al final del tratamiento (ETR), pero que experimentan una recidiva al finalizar el mismo, (pacientes: P20 y P119); en los que el RNA viral no fue detectado por PCR durante el tratamiento, pero volvió a detectarse durante los meses posteriores al tratamiento; y de 3 pacientes infectados con HCV genotipo 1a, de los cuales 2 son no respondedores (pacientes: P08 y P146), y uno es respondedor al final del tratamiento, (paciente P47).

A estos datasets les realizamos análisis de usos de codones, se determinó el Número Efectivo de Codones (ENC) (Wright, 1990) y el contenido en G+C en la 3 posición sinónima (CG3s), mediante el Programa CODON W. A continuación se determino el Uso Relativo de Codones Sinónimos (RSCU).

Las secuencias de los datasets fueron alineadas mediante MEGA 5,05 (Tamura *et al.* 2011), y se utilizó el servidor web Datamonkey (Delport *et al.* 2010), para identificar el modelo evolutivo más adecuado que mejor se ajuste a nuestros datos.

Se les realizó un análisis bayesiano, construyéndose árboles de máxima credibilidad los que fueron generados usando el programa Tree Annotator del paquete BEAST y el programa FIGTREE v1.2.2 (disponible en: <http://tree.bio.ed.ac.uk>); y un análisis de redes utilizando los enfoques de reducción de la mediana de la red, con el programa Network 4.6 (disponible en: <http://www.fluxus-engineering.com>). (Véase Materiales y Métodos).

4.1 Resultados objetivo específico 1

4.1.1 Análisis de uso de codones de pacientes respondedores, no respondedores y crónicos.

El análisis de uso de codones comparativo entre pacientes respondedores, no respondedores y crónicos mediante el Programa CODON W, permitió determinar el Uso Relativo de Codones Sinónimos (RSCU), que no reveló una utilización diferente de los mismos, como se ve en la Figura 13, los codones mas utilizados para codificar la gran mayoría de los aminoácidos, es el mismo en los pacientes indiferentemente de su respuesta a la terapia, a excepción del aminoácido Cerina (C) que es codificado mayoritariamente por el codón UGC en respondedores y no respondedores y por el codón UGU en los pacientes crónicos (Figura 13).

4.1.2. Análisis evolutivo durante el tratamiento, mediante redes y árboles de máxima credibilidad en pacientes con diferente respuesta a la terapia.

La evolución del HCV antes, durante y después del tratamiento está relacionada fundamentalmente con el hospedero. El análisis evolutivo durante todo el tratamiento mostró que cada paciente presenta diferentes dinámicas de población independiente al resultado de la terapia.

El análisis de la dinámica evolutiva de las cuasiespecies durante las primeras 4 semanas de terapia (dataset A) muestra que en el caso del paciente 7 que presenta respuesta viral sostenida, (Figura 14) existe una sustitución de secuencias durante las 4 semanas estudiadas, el pool de secuencias que se hallan en el día cero (representadas en amarillo) es reemplazado por nuevas secuencias a través de los distintos tiempos, lo cual se ve claramente tanto en las redes (Figura 14 c) como en el árbol de máxima credibilidad (Figura 14 b), quedando las secuencias circulantes en la cuarta semana (representadas en verde) mas emparentadas con las circulantes en las semanas anteriores que con las que se hallaban presentes al comienzo de la terapia. A su vez se destaca que las secuencias presentes al inicio de la terapia guardan bastante homogeneidad.

En cambio en el caso del paciente 9 que también presenta respuesta viral sostenida, (Figura 15) existen secuencias que se encuentran presentes durante toda la terapia, en la red se visualizan dichos haplotipos asociados a los 4 muestreos identificados por los 4 colores, (Figura 15 c).

La dinámica evolutiva de las cuasiespecies en el caso del paciente 10, que no responde a la terapia, mediante el análisis por redes o por árboles de máxima credibilidad muestra que las secuencias presentes en la cuarta semana de terapia están en su mayoría más emparentadas con las secuencias presentes al inicio de la terapia que con las que aparecen en tiempos posteriores (Figura 16).

En cambio en el caso del paciente 12 que también es no respondedor (Figura 17) existen secuencias que se encuentran presentes durante toda la terapia, en la red se visualizan dichos haplotipos asociados a los 4 muestreos identificados por sus respectivos colores.

Por su parte, en el paciente 16, que es un paciente crónico sin tratamiento (Figura 18), algunas de las secuencias presentes al cabo de 10 meses se hallan asociadas a secuencias presentes al inicio del estudio y otras se relacionan con secuencias existentes en períodos de tiempo anteriores.

4.1.3 Análisis evolutivo antes, durante y posterior al tratamiento, mediante redes y árboles de máxima credibilidad en pacientes con diferente respuesta a la terapia y de diferente genotipo.

El análisis evolutivo llevado a cabo antes, durante y durante 5 meses posteriores de haber finalizado el tratamiento mostró que en cada paciente se presentan diferentes dinámicas de población, independiente al resultado de la terapia y al genotipo perteneciente.

En el análisis la dinámica evolutiva de las cuasiespecies durante este período (dataset B) muestra que en el caso del paciente P75, (Figura 19) que es un paciente no respondedor del genotipo 3a, la mayoría de las secuencias que constituyen la nube de la cuasiespecie circulante a los 5 meses pos-terapia se hallaban mas emparentadas con secuencias presentes en la pre-terapia que con las secuencias que surgieron durante la terapia.

A su vez en el paciente P145, que es un paciente no respondedor del genotipo 3a (Figura 20), las secuencias que constituyen la nube de la cuasiespecie circulante a los 5 meses pos-terapia se hallan emparentadas con secuencias que surgieron durante la terapia, y no con las que se encontraban en la pre-terapia.

El paciente P08, perteneciente al genotipo 1a, no respondedor (Figura 21) las secuencias obtenidas al mes 6 pos-terapia se hallan asociadas a secuencias que surgieron con posterioridad de la terapia.

En el caso del paciente P07, otro paciente no respondedor del genotipo 3a (Figura 22), la mayoría de las secuencias encontradas circulando al quinto mes pos-terapia se encontró que estaban asociadas a secuencias que surgieron durante la terapia, aunque también circularon otras secuencias mas emparentadas con las presentes en la pre-terapia.

Por su parte el paciente P146, paciente no respondedor del genotipo 1a (Figura 23) todas las secuencias obtenidas al mes 6 pos-terapia se hallan asociadas a secuencias que estaban presentes en la pre-terapia.

Por su parte, el paciente P119, que es un paciente respondedor al final de la terapia del genotipo 3a (Figura 24), las secuencias encontradas 5 meses pos-terapia no se hallan asociadas a secuencias presentes en la pre-terapia, sino que se asocian a secuencias presentes en períodos anteriores con presencia de eventos homoplásticos visibles en el árbol de máxima verosimilitud.

En el caso del paciente P20, un paciente respondedor al final del tratamiento, perteneciente al genotipo 3a (Figura 25), casi todas las secuencias presentes encontradas 5 meses pos-terapia se hallan asociadas a secuencias presentes en la pre-terapia.

Mientras que los análisis realizados al paciente P47, un paciente respondedor al final del tratamiento, perteneciente al genotipo 1a (Figura 26), las secuencias circulantes a los 6 meses pos-tratamiento estaban en su mayoría asociadas a secuencias que surgieron inmediatamente después de finalizado el tratamiento, y unas pocas se hallan asociadas a secuencias presentes en periodos de tiempo anteriores.

4.1.4. Análisis coalescente bayesiano de las secuencias de la NS5A de cuasiespecies circulantes durante durante los tratamientos en los diferentes pacientes.

Los análisis de los skyline resultantes de los análisis bayesianos muestra que el tamaño efectivo de la población viral de los pacientes mostró un fuerte descenso al inicio de la terapia, y que en la mayoría de los casos recupero se variabilidad con el transcurrir del tiempo.

En el caso de los pacientes crónicos el Skyline no mostró descenso dado que no existió presión inmune por no haber tratamiento.

Codon	Pacientes SR			Pacientes NR			Pacientes C		
	P7	P8	P9	P10	P11	P12	P16	P17	P18
UUU(F)	0,83	0	0,83	0,27	0,02	0,62	0,26	0,68	0
UUC(F)	1,17	2	1,17	1,73	1,98	1,38	1,74	1,32	2
UUA(L)	0,01	0,25	0,01	0,02	0,24	0,25	0,01	0,07	0,24
UUG(L)	1,17	1,25	1,09	0,95	0,79	1,46	1,11	0,51	1,43
CUU(L)	0,96	0,75	0,8	0,87	0,97	0,51	1,1	0,79	0,28
CUC(L)	1,63	1,75	1,93	1,82	1,54	2,1	1,62	1,82	2,08
CUA(L)	0,05	0,25	0,01	0,6	0,26	0,02	0,81	0,74	0,24
CUG(L)	2,19	1,75	2,16	1,74	2,2	1,65	1,34	2,06	1,73
AUU(I)	0,72	0,43	1,14	0,37	0,85	0,82	0,44	0,73	0,67
AUC(I)	1,09	1,72	0,78	1,09	1,64	1,07	1,26	0,72	1,05
AUA(I)	1,19	0,86	1,08	1,54	0,51	1,1	1,31	1,55	1,28
GUU(V)	0,97	0,92	1,17	1,03	1,05	0,98	1,01	0,73	0,95
GUC(V)	0,8	0,76	0,93	0,98	0,61	0,77	0,89	1	0,93
GUA(V)	0,82	0,92	0,85	0,94	1,43	0,51	1,02	0,92	0,96
GUG(V)	1,4	1,39	1,04	1,05	0,92	1,74	1,09	1,35	1,16
UCU(S)	1,41	1,16	1,35	1,6	1,79	1	1,71	1,37	1,21
UCC(S)	2,58	2,7	2,52	2,14	2,38	2,64	2,1	2,47	2,44
UCA(S)	0,59	0,51	0,58	0,72	0,48	0,55	0,68	0,61	0,68
UCG(S)	0,43	0,65	0,58	0,57	0,52	0,72	0,56	0,59	0,71
AGU(S)	0	0	0	0,2	0,2	0,18	0,18	0,21	0,19
AGC(S)	0,98	0,97	0,97	0,78	0,62	0,91	0,77	0,75	0,77
CCU(P)	0,94	0,61	0,82	1,15	0,92	0,73	0,97	0,98	0,8
CCC(P)	1,42	1,46	1,3	1,07	1,2	1,44	1,28	1,26	1,26
CCA(P)	0,81	0,96	0,64	0,89	0,94	0,84	0,78	0,84	0,92
CCG(P)	0,84	0,97	1,24	0,89	0,93	0,99	0,97	0,92	1,02
ACU(T)	0,7	0,74	0,54	0,36	0,86	0,45	0,52	0,46	0,59
ACC(T)	1,2	1,24	1,64	1,64	1,28	1,55	1,34	1,57	1,27
ACA(T)	1,43	1,34	0,89	1,32	1,07	1,13	1,42	1,16	1,23
ACG(T)	0,66	0,68	0,93	0,67	0,79	0,87	0,72	0,81	0,91
GCU(A)	0,85	0,8	0,39	0,8	0,51	0,95	0,59	1,08	1,07
GCC(A)	1,48	1,38	1,7	1,44	1,42	1,35	1,65	1,27	1,01
GCA(A)	0,55	0,55	0,35	0,22	0,91	0,2	0,57	0,19	0,99
GCG(A)	1,12	1,26	1,57	1,54	1,16	1,5	1,19	1,46	0,93
UAU(Y)	0,01	0,01	0	0,42	0,02	0,04	0,63	0,52	0,53
UAC(Y)	1,99	1,99	2	1,58	1,98	1,96	1,37	1,48	1,47
CAU(H)	1,37	0,65	1	0,62	0,7	0,53	0,89	0,06	0,66
CAC(H)	0,63	1,35	1	1,38	1,3	1,47	1,11	1,94	1,34
CAA(Q)	0,02	0	0,41	0,55	0,34	0,8	0,4	0,45	0,42
CAG(Q)	1,98	2	1,59	1,45	1,66	1,2	1,6	1,55	1,58
AAU(N)	0,76	0,12	0,4	0,45	0,27	0,36	0,39	0,31	0,33
AAC(N)	1,24	1,88	1,6	1,55	1,73	1,64	1,61	1,69	1,67
AAA(K)	0,45	0,67	0,3	0,7	0,4	0,43	0,88	0,53	0,92
AAG(K)	1,55	1,33	1,7	1,3	1,6	1,57	1,12	1,47	1,08
GAU(D)	0,34	0,11	0,3	0,59	0,42	0,43	0,45	0,66	0,38
GAC(D)	1,66	1,89	1,7	1,41	1,58	1,57	1,55	1,34	1,62
GAA(E)	0,53	0,43	0,56	0,5	0,31	0,44	0,43	0,38	0,47
GAG(E)	1,47	1,57	1,44	1,5	1,69	1,56	1,57	1,62	1,53
UGU(C)	0,66	0,05	0,22	1,3	0,01	0,64	1,07	1,27	1,22
UGC(C)	1,34	1,95	1,78	0,7	1,99	1,36	0,93	0,73	0,78
CGU(R)	0,53	0,73	0,5	0,86	0,87	0,92	0,64	0,72	0,64
CGC(R)	0,67	1,46	0,49	0,86	0,88	0,48	0,89	0,58	1,2
CGA(R)	0,44	0,37	0,53	0,76	0,46	0,44	0,84	0,42	0,15
CGG(R)	1,69	0,73	1,95	1,25	1,32	1,39	1,24	1,67	0,82
AGA(R)	0,37	0	0,52	0,43	0,62	0,46	0,41	0,62	1,23
AGG(R)	2,3	2,7	2,01	1,83	1,84	2,32	1,98	1,99	1,97
GGU(G)	0,31	0,01	0	0,03	0,31	0,34	0	0,03	0,01
GGC(G)	0,94	1,64	1,15	1,51	0,94	0,66	1,53	1,48	1,33
GGA(G)	0,63	0,5	0,57	0,84	0,92	0,04	0,9	0,89	0,01
GGG(G)	2,12	1,85	2,27	1,62	1,84	2,95	1,57	1,6	2,65

Figura 13. Uso comparativo de codones entre los pacientes no respondedores NR, pacientes con respuesta sostenida RS y pacientes crónicos C, correspondientes al dataset A. En negrita y rojo se marca el codón mas utilizado para cada aminoácido.

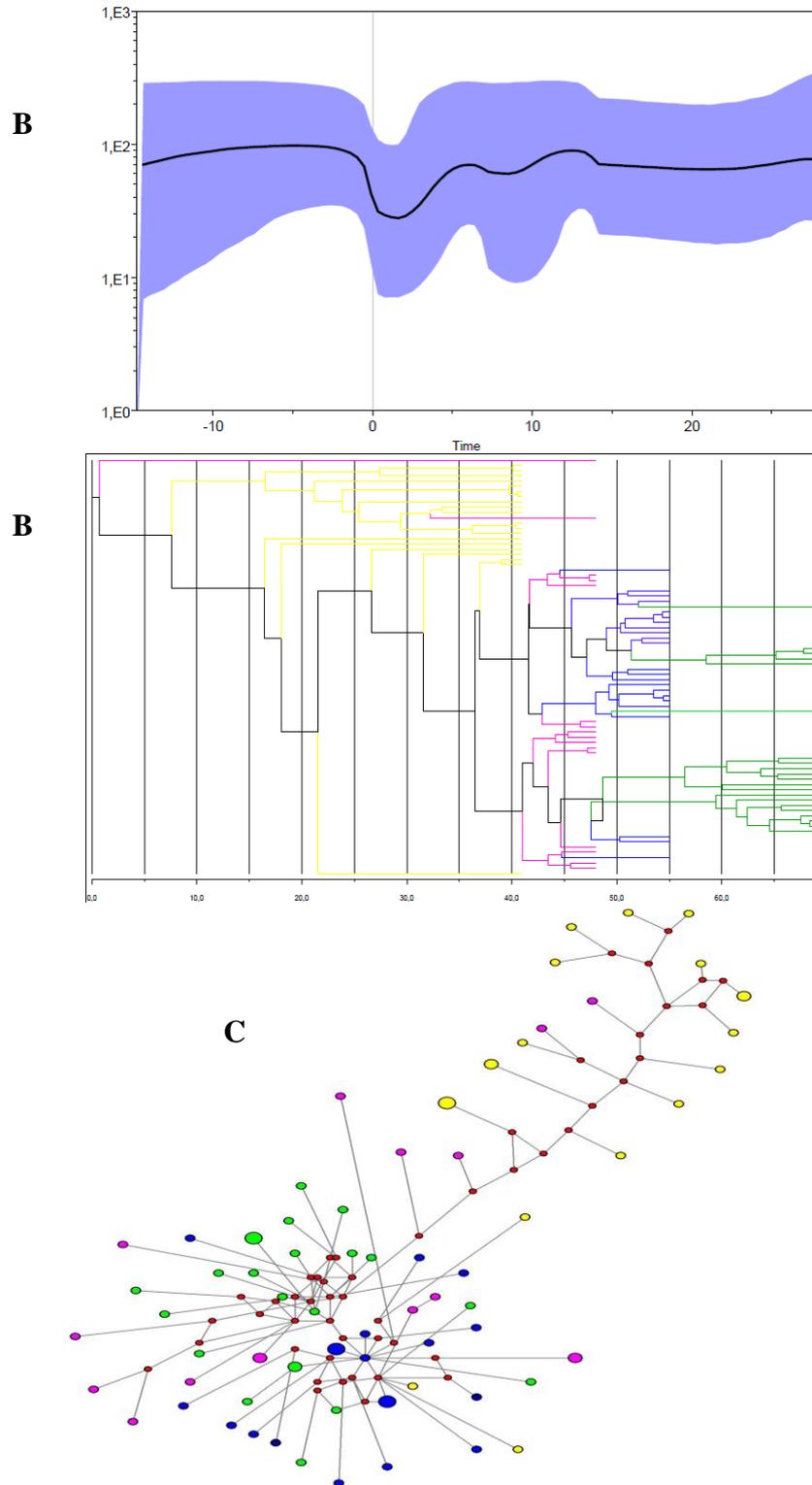


Figura 14. Skyline en días, la línea vertical indica el comienzo de la terapia (A), Arbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente 7 con respuesta viral sostenida. Los colores en (B) y (C) representan: amarillo día 0, lila día 7, azul día 14 y verde día 28 de la terapia; en rojo se representan los vectores medios de la red. Los bordes representan diferentes haplotipos y su tamaño representa el número de haplotipos presentes.

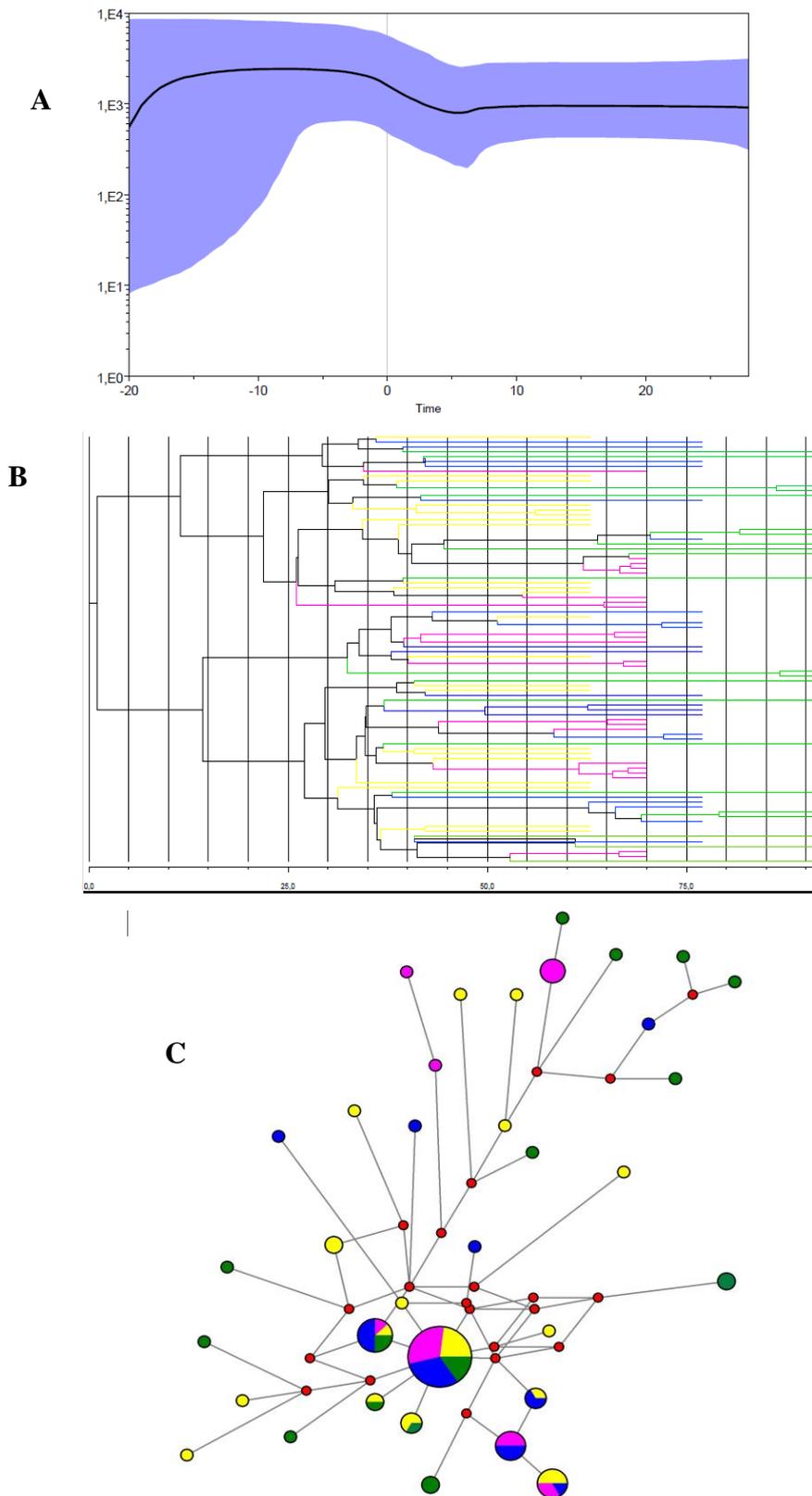


Figura 15. Skyline en días, la línea vertical indica el comienzo de la terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente 9 con respuesta viral sostenida. Los colores en (B) y (C) representan: amarillo día 0, lila día 7, azul día 14 y verde día 28 de la terapia; en rojo se representan los vectores medios de la red.

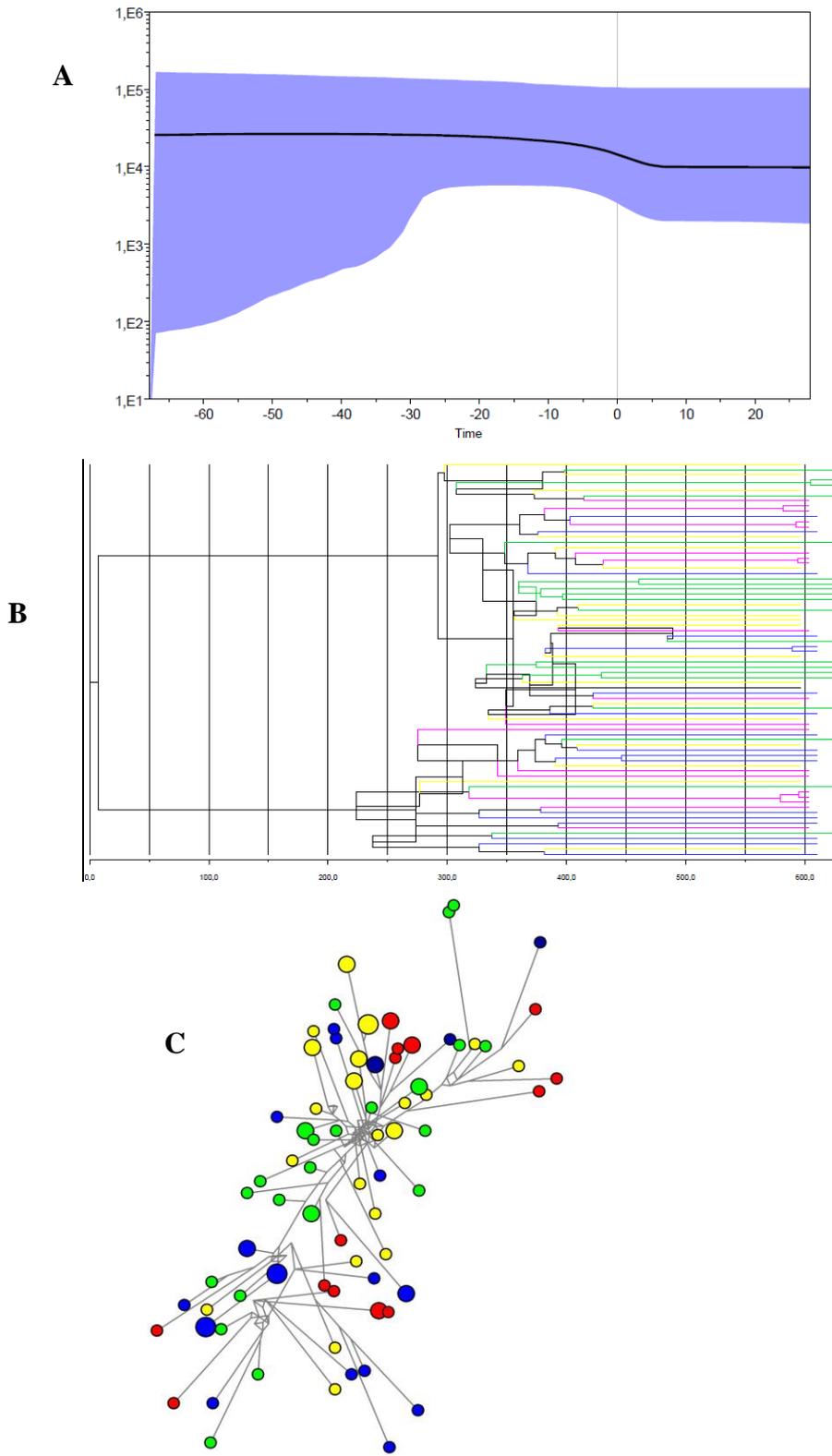


Figura 16. Skyline en días, la línea vertical indica el comienzo de la terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente 10 no respondedor. Los colores en (B) y (C) representan: amarillo día 0, rojo día 7, azul día 14 y verde día 28 de la terapia. Los bordes representan diferentes haplotipos y su tamaño representa el número de haplotipos presentes.

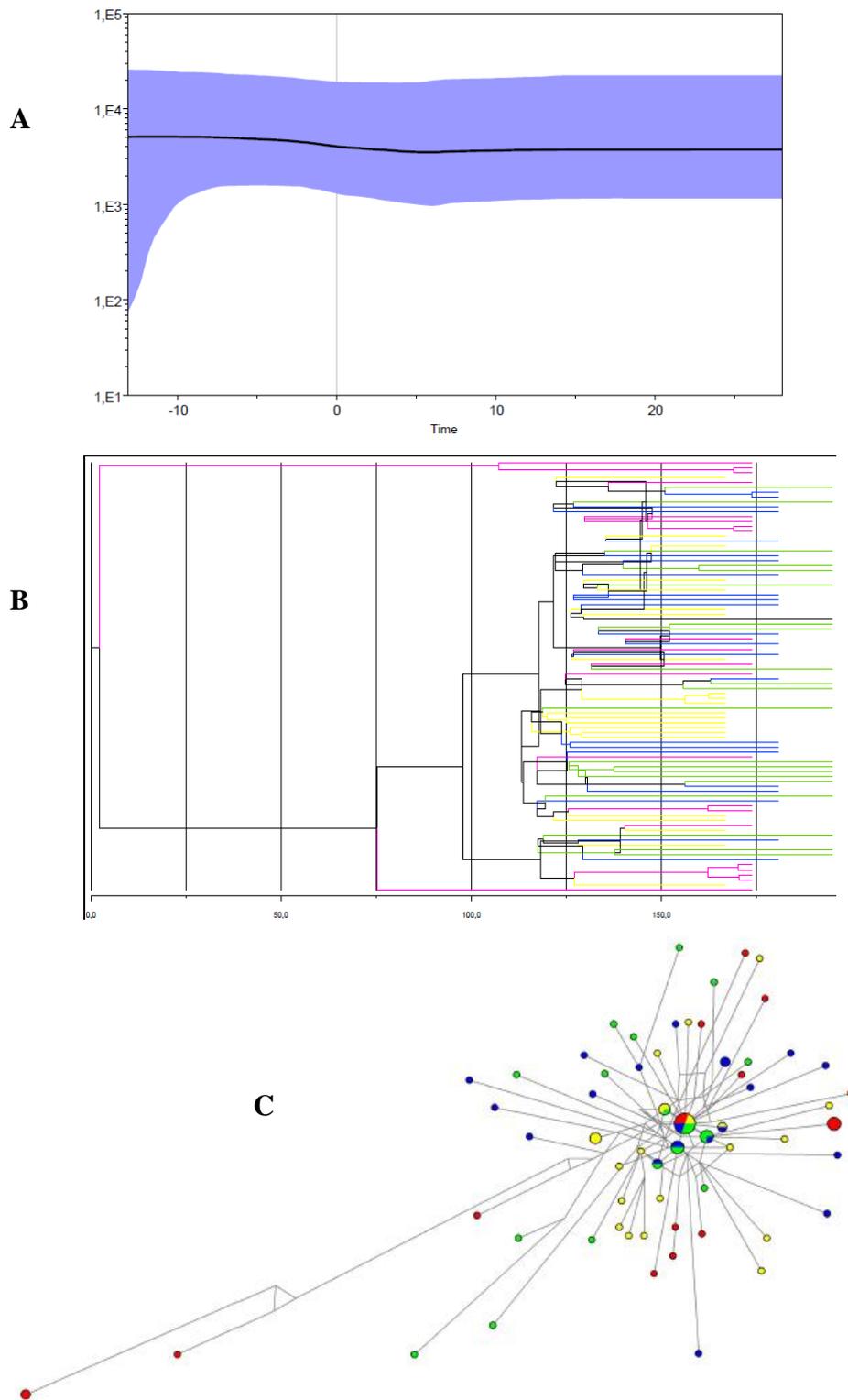


Figura 17. Skyline en días, la línea vertical indica el comienzo de la terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente 12 no respondedor. Los colores en (B) y (C) representan: amarillo día 0, rojo día 7, azul día 14 y verde día 28 de la terapia. Los bordes representan diferentes haplotipos y su tamaño representa el número de haplotipos presentes.

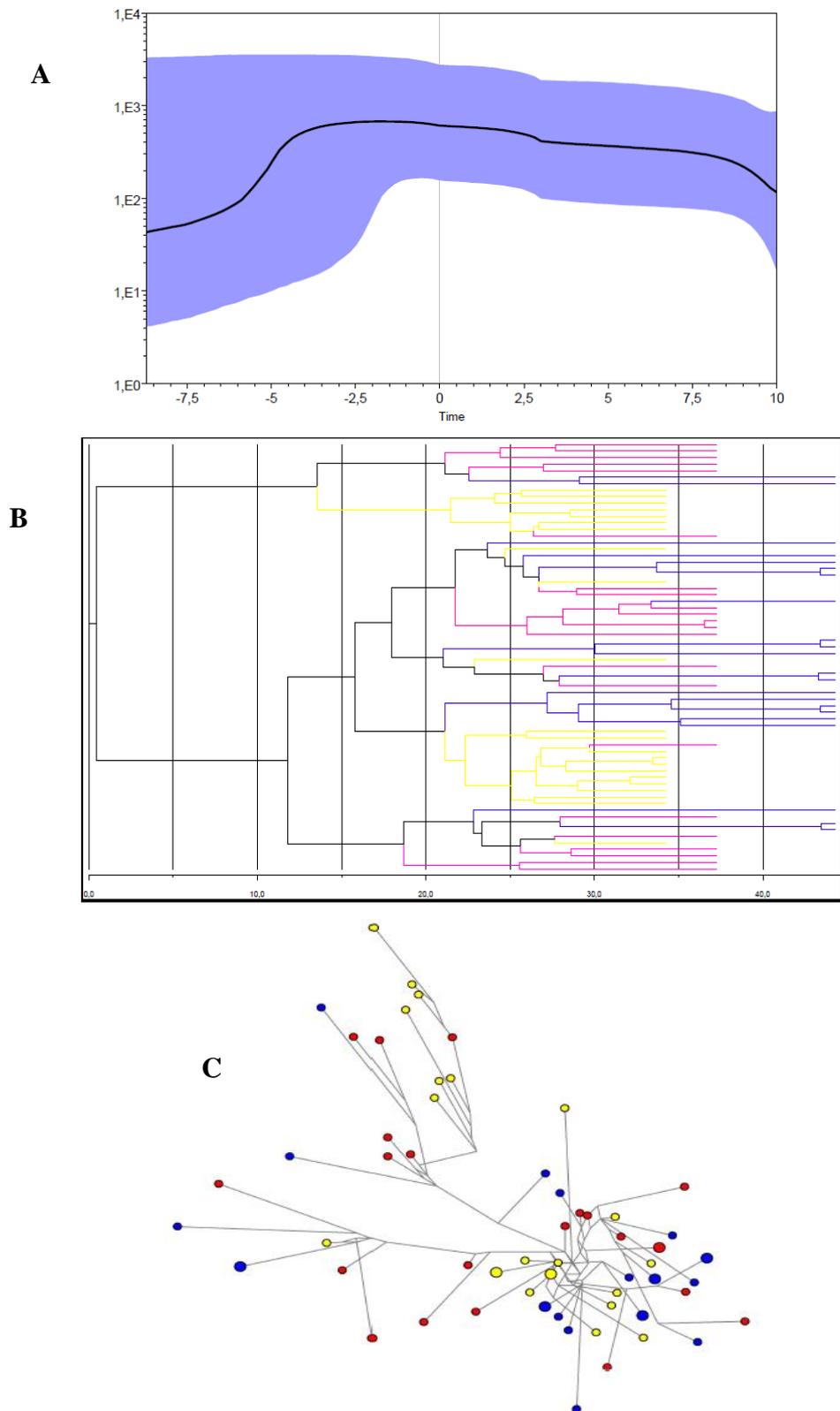


Figura 18. Skyline en meses, la línea vertical indica el comienzo de las tomas (A), Arbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente 16 crónico sin tratamiento. Los colores en (B) y (C) representan: amarillo mes 0, rojo mes 3, azul mes 10 a partir de la primera toma. Los bordes representan diferentes haplotipos y su tamaño representa el número de haplotipos presentes.

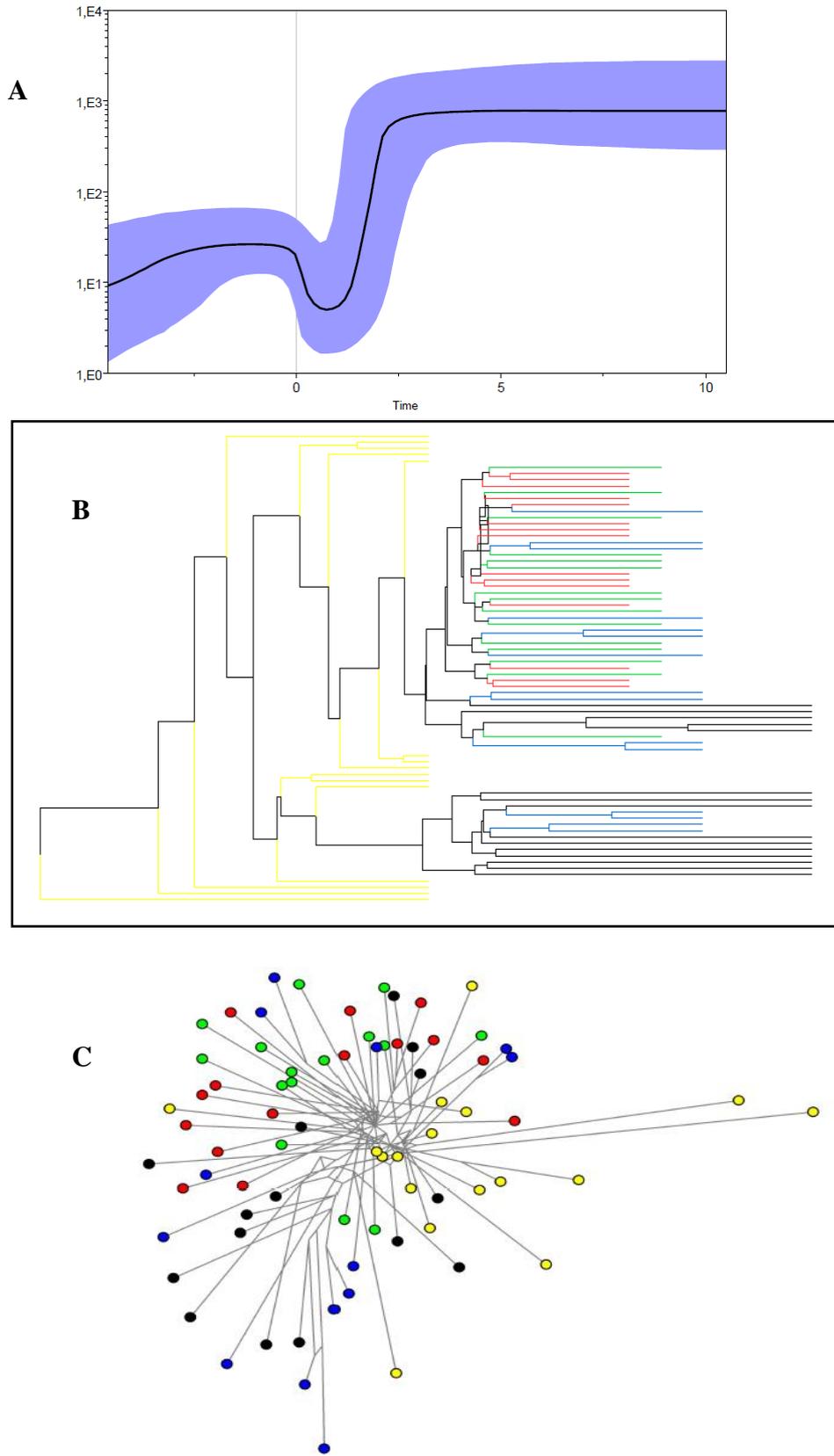


Figura 19. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P75, no respondedor del genotipo 3a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo tratamiento, verde 21 días, azul 2 meses y negro 5 meses pos-tratamiento.

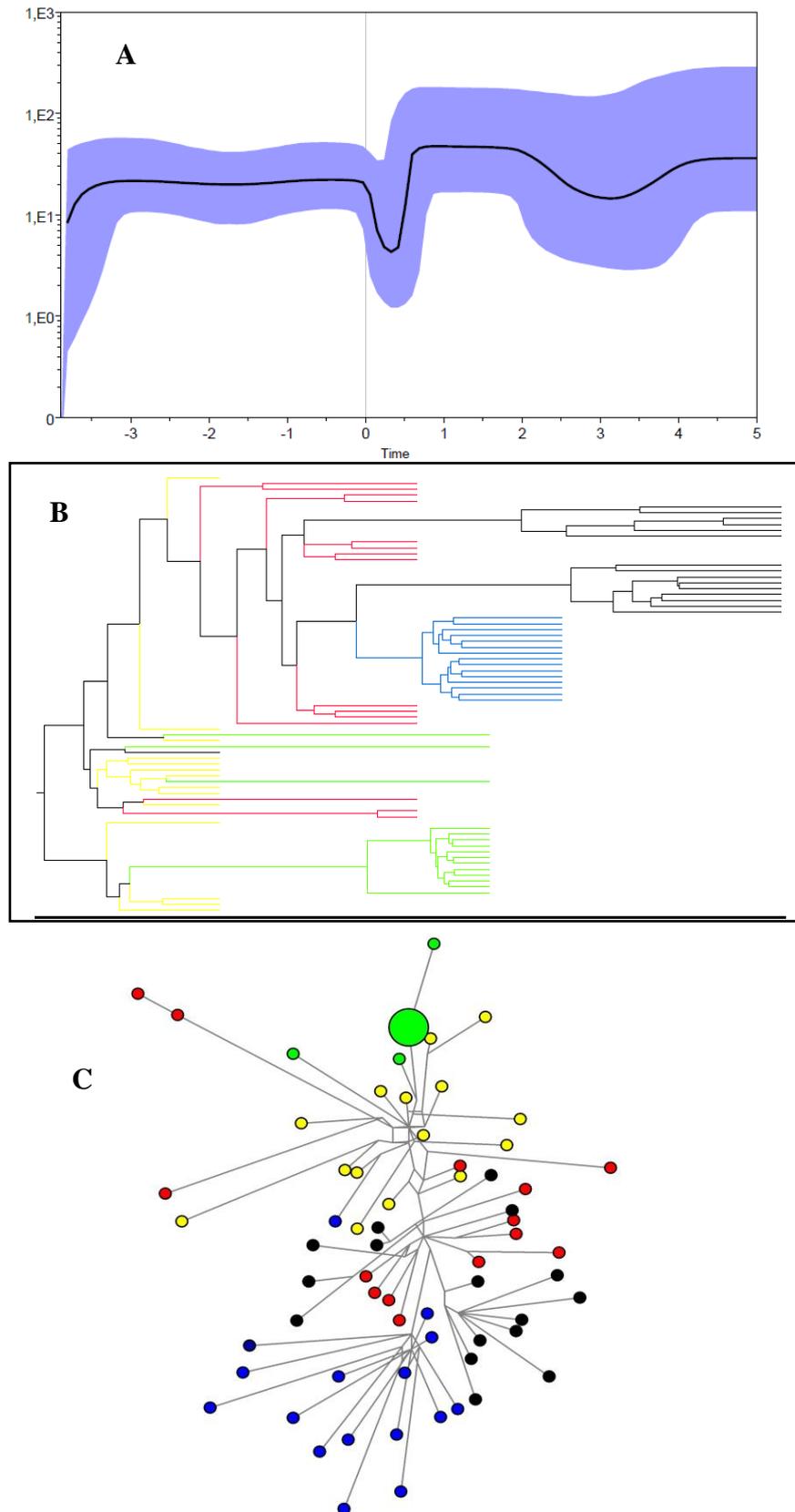


Figura 20. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P145, no respondedor del genotipo 3a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo tratamiento, verde 28 días, azul 2 meses y negro 5 meses pos-tratamiento.

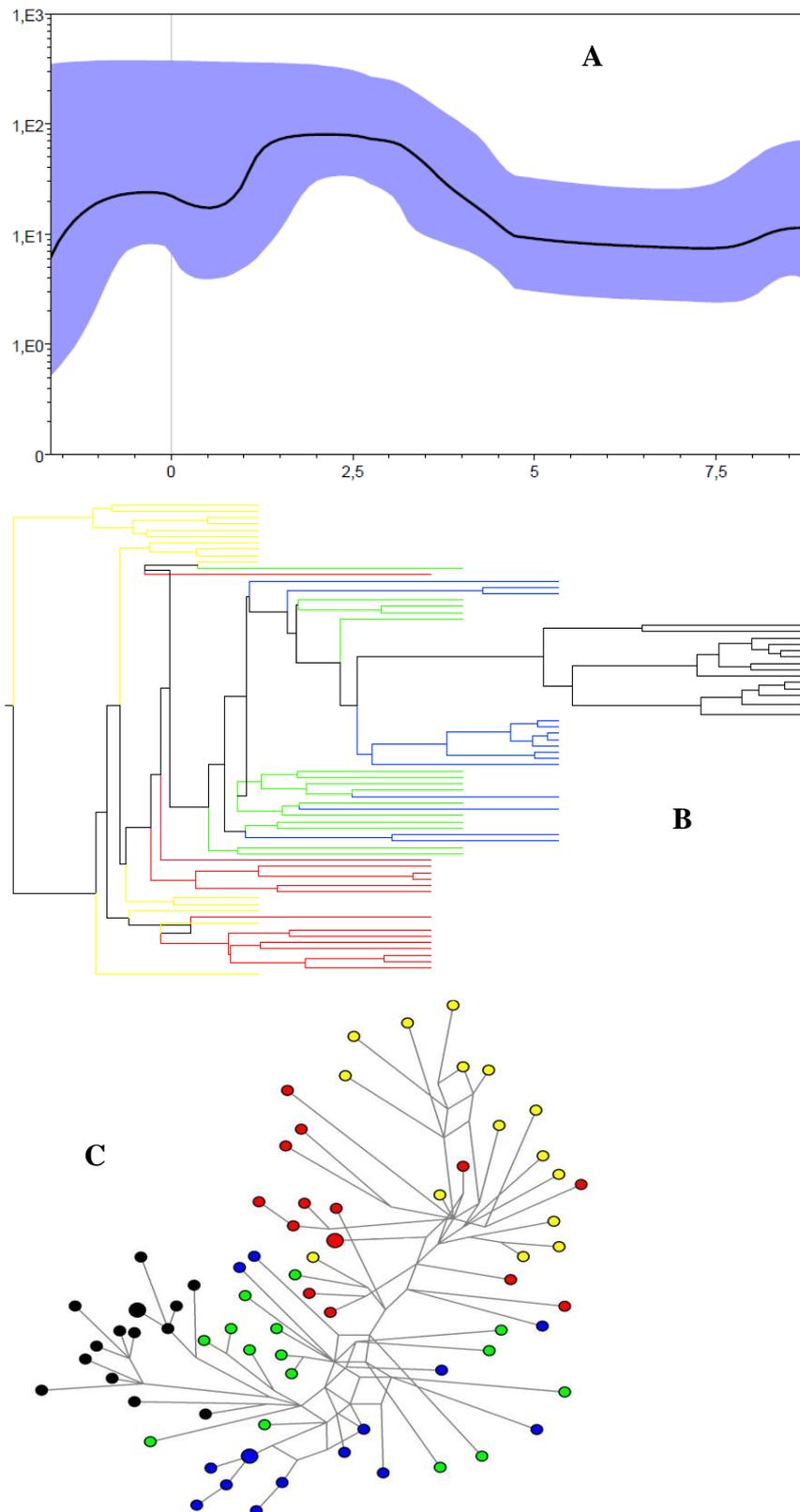


Figura 21. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Arbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P08, no respondedor del genotipo 1a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo tratamiento, verde 14 días pos-tratamiento, azul 2 meses pos-tratamiento y negro 6 meses pos-tratamiento.

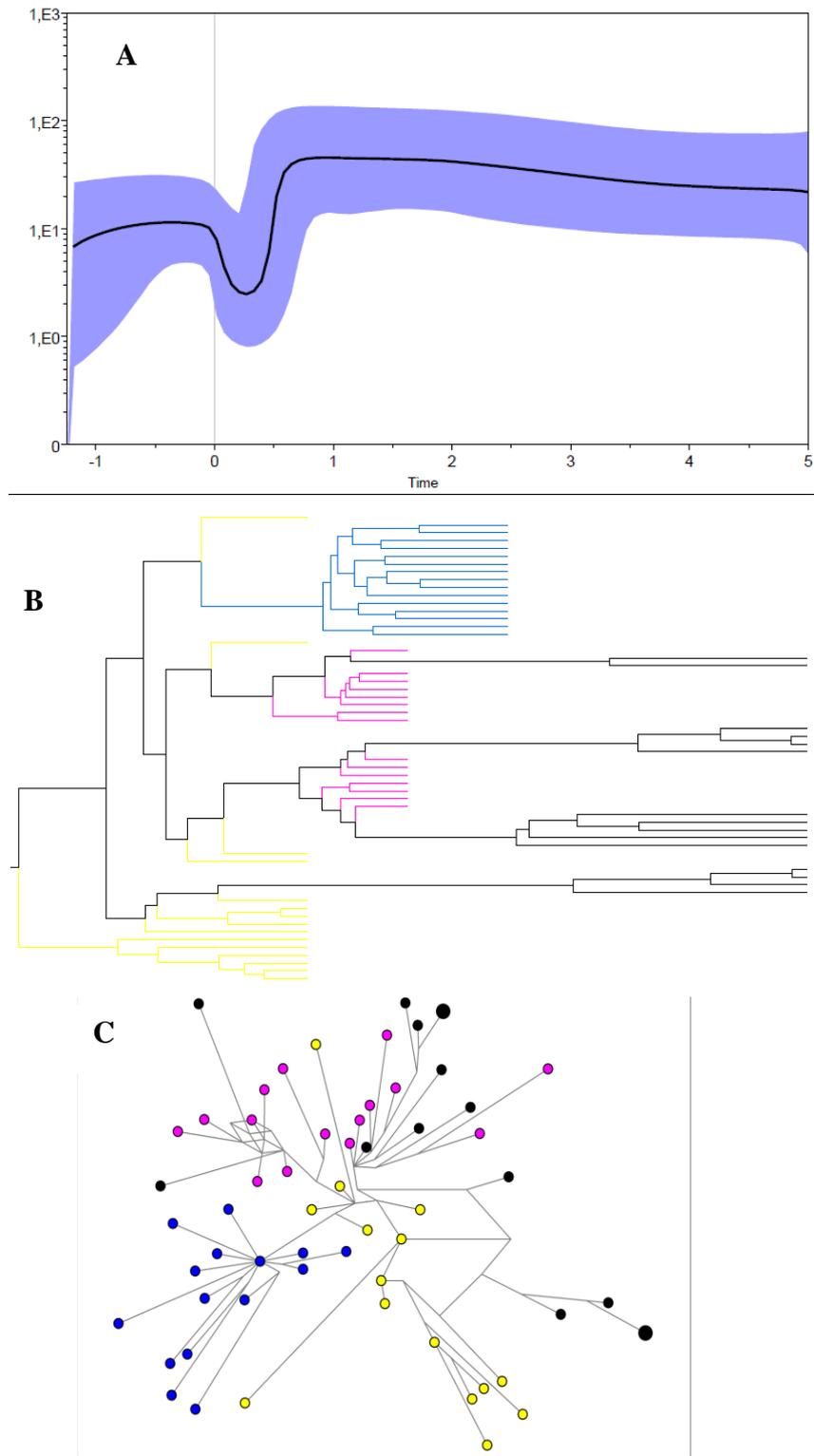


Figura 22. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P07, no respondedor del genotipo 3a. Los colores en (B) y (C) representan: amarillo pre-terapia, lila 28 días pos-tratamiento, azul 4 meses pos-tratamiento y negro 5 meses pos-tratamiento.

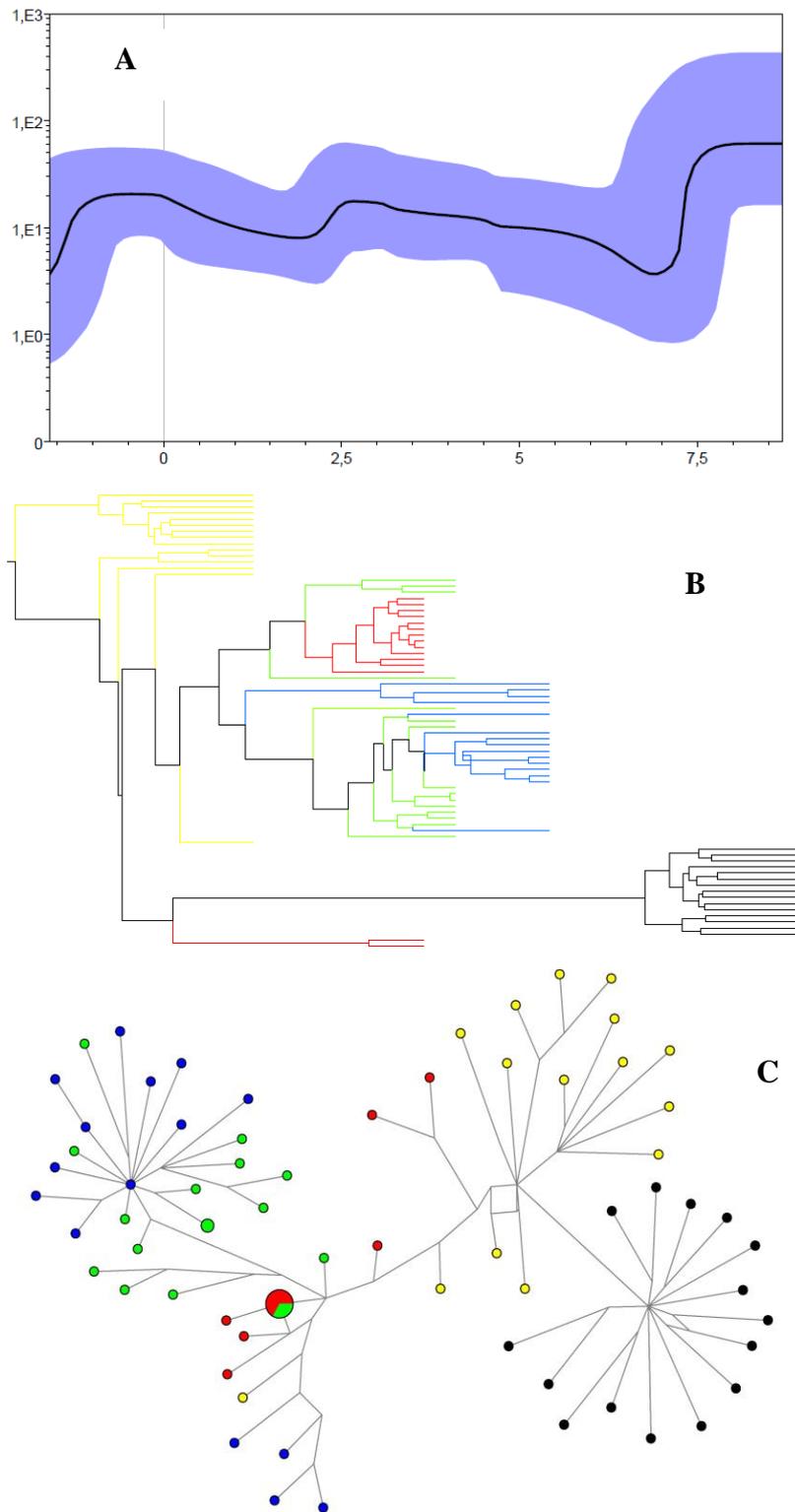


Figura 23. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P146, no respondedor del genotipo 1a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo tratamiento, verde 14 días pos-tratamiento, azul 2 meses pos-tratamiento, negro 6 meses pos-tratamiento.

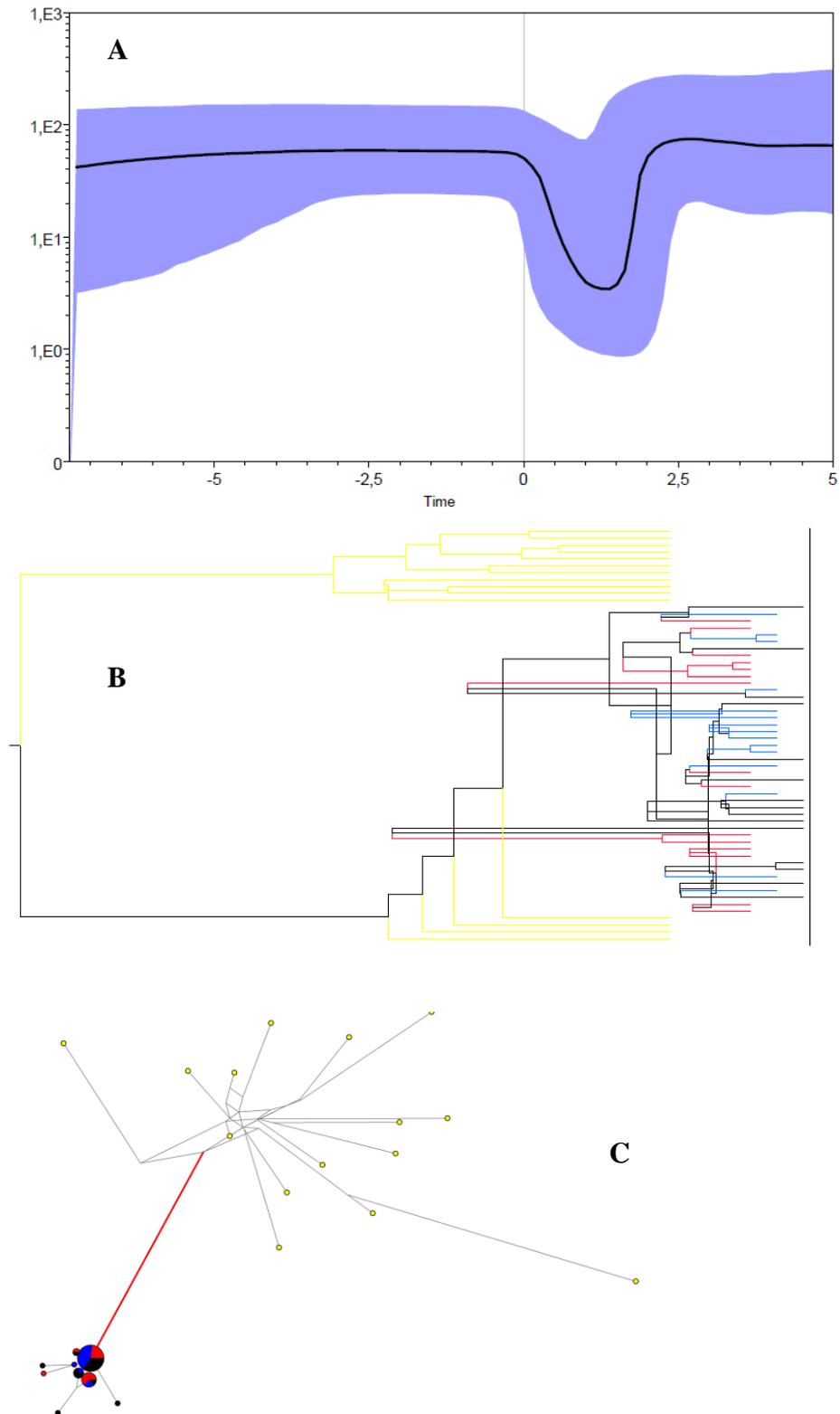


Figura 24. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P119, respondedor al final del tratamiento del genotipo 3a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo 3 meses pos-tratamiento, azul 4 meses pos-tratamiento, y negro 5 meses pos-tratamiento.

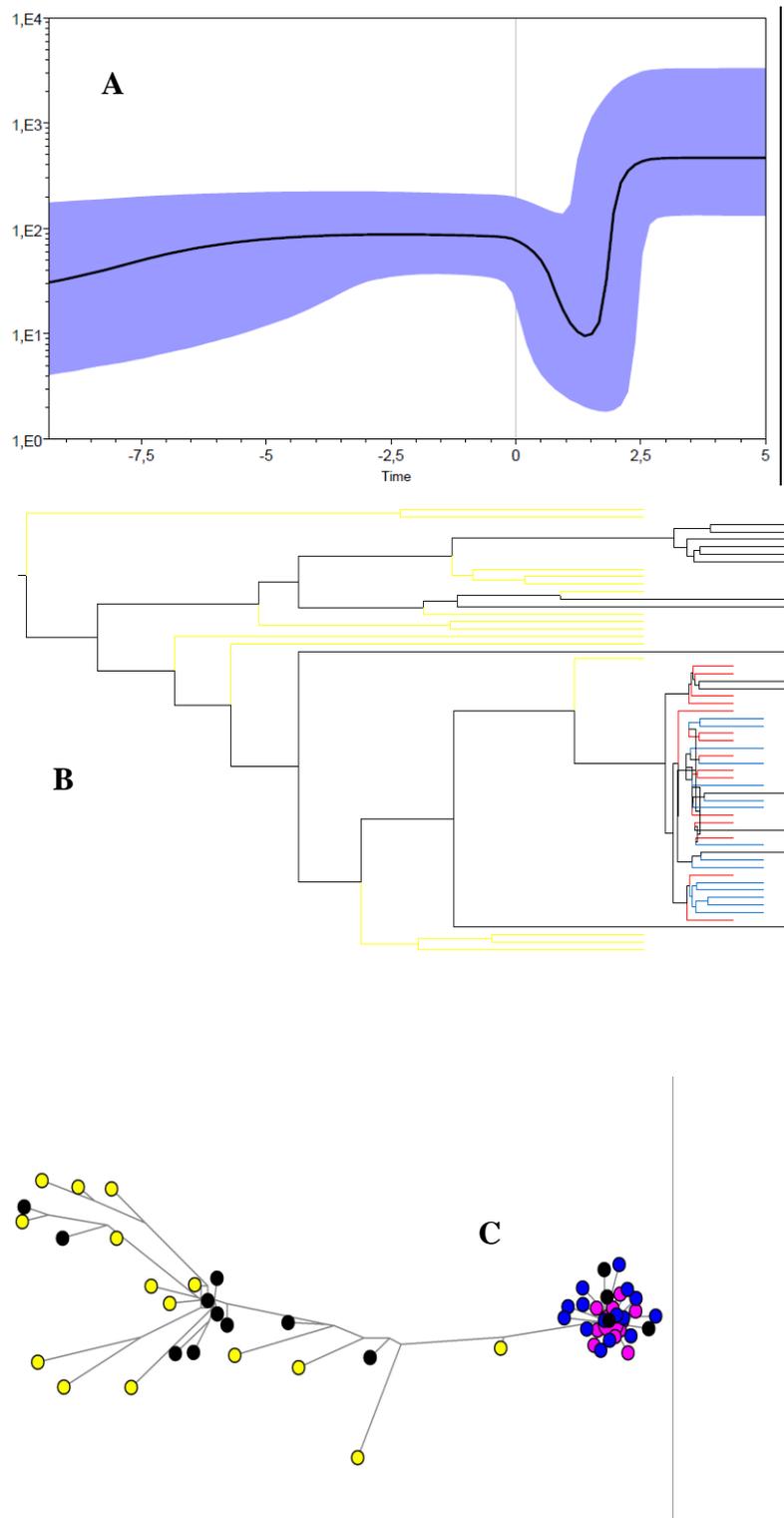


Figura 25. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P20, respondedor al final del tratamiento del genotipo 3a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo 3 meses pos-tratamiento, azul 4 meses pos-tratamiento, y negro 5 meses pos-tratamiento.

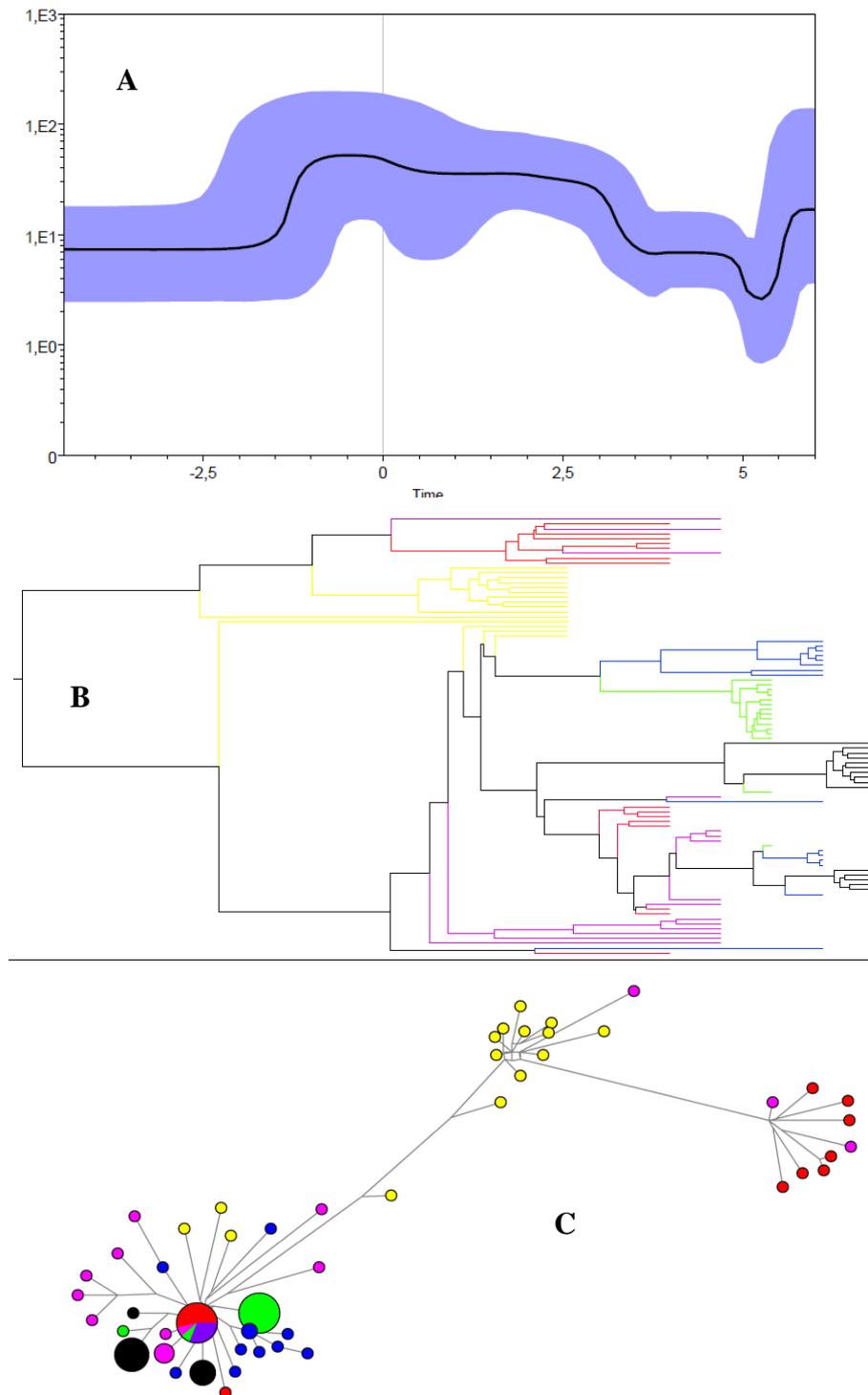


Figura 26. Skyline en meses, la línea vertical indica el comienzo de las terapia (A), Árbol de máxima credibilidad (B), y Redes (C) correspondientes al paciente P47, respondedor al final del tratamiento del genotipo 3a. Los colores en (B) y (C) representan: amarillo pre-terapia, rojo 2 meses pos-tratamiento, lila 3 meses pos-tratamiento, verde 4 meses pos-tratamiento, azul 5 meses pos-tratamiento y negro 6 meses pos-tratamiento.

Tabla 1. Resume de las relaciones obtenidas de las secuencias de la cuasiespecie al final de la terapia con las circulantes antes o durante el tratamiento. En los pacientes (P75, P146, P20) las secuencias en los tiempos finales se asocian a las secuencias en la pre-terapia, mientras que en los pacientes (P145, P08, P119, P47) dichas secuencias se asocian a secuencias que surgieron durante el tratamiento, en otros pacientes (P07) las mismas secuencias se hallaron durante todo el período.

Paciente	Respuesta	Asociación con secuencias durante la terapia
7	RS	Semanas anteriores
9	SR	Presentes durante toda la terapia
10	NR	Presentes al inicio de la terapia
12	NR	Presentes durante toda la terapia
16	C	Inicio de la terapia y anteriores en el tiempo

Paciente	Genotipo	Respuesta	Asociación con secuencias antes y durante la terapia
75	3a	NR	Asociadas a secuencias en pre-terapia
145	3a	NR	Asociadas a secuencias durante la terapia
7	3a	NR	Durante (mayoría) Pre-terapia (minoría)
8	1a	NR	Asociadas a secuencias posteriores a la terapia
146	1a	NR	Asociadas a secuencias en pre-terapia
119	3a	ETR	Asociadas a secuencias durante la terapia
20	3a	ETR	Asociadas a secuencias en pre-terapia
47	1a	ETR	Asociadas a secuencias finalizado el tratamiento

RS. -Respuesta viral sostenida

NR. - No Respondedor

ETR. -Respondedor al final del tratamiento

4.2 Discusión del objetivo específico 1

El HCV es un virus muy variable, lo que le confiere un abanico de posibilidades para que pueda evolucionar y adaptarse a las nuevas condiciones.

4.2.1 Análisis de uso de codones de pacientes respondedores, no respondedores y crónicos.

El análisis de RSCU, correspondientes a la región de la NS5A del HCV, (nt 6786 a 7547) entre pacientes respondedores, no respondedores y crónicos del genotipo 1a no reveló una utilización diferente de los mismos, (Figura 13) Los codones mas utilizados para codificar la gran mayoría de los aminoácidos es el mismo en los pacientes indiferentemente de su respuesta a la terapia, a excepción del aminoácido cerina (C) que es codificado mayoritariamente por el codón UGC en respondedores y no respondedores y por el codón UGU en los pacientes crónicos.

Esta similitud en los RSCU tal vez debido a lo corto de la secuencia analizada (762 nucleótidos, correspondiente a 254 codones).

4.2.2 Análisis evolutivo durante el tratamiento, mediante redes y árboles de máxima credibilidad en pacientes con diferente respuesta a la terapia.

Si bien se observan patrones evolutivos diferentes en diferentes pacientes, esto es claramente visualizado mediante el estudio de redes complejas.

Las redes reticuladas proporcionan una visión explícita de la evolución. En este tipo de red, los bordes representan linajes de descendencia, en nuestro caso haplotipos, y los nodos corresponden a los ancestros hipotéticos; el tamaño de los haplotipos representa la cantidad de secuencias presentes del mismo.

La implementación de análisis de redes, como complemento de los análisis filogenéticos realizado mediante la construcción de árboles de máxima verosimilitud, con métodos bayesianos, nos permitió una forma mas clara de visualizar las variaciones evolutivas experimentadas por la nube de mutantes dentro de la cuasiespecie a través del tiempo, nos permite identificar que secuencias permanecen presentes en los distintos tiempos, así como sus relaciones con las que se hallaban antes o después

temporalmente. Nos permite visualizar mejor la homogeneidad o heterogeneidad de la población en cada etapa y como esta varia a lo largo del tiempo.

Las redes filogenéticas podrían pues desempeñar un importante papel en la reconstrucción de la historia evolutiva, dado que los abordajes con redes complejas permiten visualizar mejor las relaciones entre los haplotipos circulantes en la nube de las cuasiespecies y facilitan el análisis de los modelos evolutivos implicados.

Es por ello que consideramos a los análisis de redes como un elemento complementario de gran importancia para el estudio de las cuasiespecies.

El análisis del comportamiento de la evolución durante todo el tratamiento mostró que cada paciente presenta diferentes dinámicas de población ajeno a resultado de la terapia y al genotipo perteneciente.

La tabla 1 resume las relaciones obtenidas de las secuencias de la cuasiespecie al final de la terapia con las circulantes antes o durante el tratamiento, lo que nos permito identificar 3 tipos de pacientes, aquellos en los cuales las secuencias en los tiempos finales se asocian a las secuencias en la pre-terapia, (P75, P146, P20), otros en los que dichas secuencias se asocian a secuencias que surgieron durante el tratamiento, (P145, P08, P119, P47) y otros en los que las mismas secuencias se hallaron durante todo el período (P07); independientemente de la respuesta al tratamiento, y al genotipo perteneciente.

Lo que sugiere que algunos pacientes presentaban antes de la terapia una población viral que persiste después del tratamiento, lo que sugiere que estaban bien adaptados para evadir el tratamiento, en otros en cambio fue la presión selectiva originada por el tratamiento lo que dio lugar a la aparición de cepas resistentes capaces de evadir al tratamiento.

El hecho de que dichas cepas no son detectadas durante el tratamiento sugiere la importancia de los genomas memorias, es decir variantes capaces de replicar en bajas tasas, pero de persistir y volver a resurgir una vez que la presión selectiva deja de actuar.

Es de suponer que, con el fin de evadir la terapia, el virus replica en bajas tasas y en algún momento produce cepas con una mejor aptitud a la nueva condición, capaces de eludir el tratamiento, las que se seleccionan, constituyendo un nuevo pool de secuencias mejor adaptadas que resulta en la recaída, o resurge la cepa original una vez finalizada la presión ejercida por la terapia.

4.2.3 Análisis coalescente bayesiano de las secuencias de la NS5A de cuasiespecies circulantes durante los tratamientos en los diferentes pacientes

Los análisis de los skyline resultantes de los análisis bayesianos muestra que el tamaño efectivo de la población viral de los pacientes mostró un fuerte descenso al inicio de la terapia, y que en la mayoría de los casos recupero se variabilidad con el transcurrir del tiempo. Al inicio del tratamiento la presión selectiva ejercida sobre la cuasiespecie genera un descenso en la variabilidad, pero en algún punto del tiempo durante el tratamiento surge una cepa predominante, o resurge una cepa que en una baja frecuencia permaneció como genoma memoria en la población de la cuasiespecie, con lo que la población viral con el tiempo recupera su variabilidad.

5. Objetivo específico 2

Analizar dinámica evolutiva de las cuasiespecies en distintas regiones del genoma en un mismo paciente crónico sin terapia.

La variación genética observada a lo largo del genoma de HCV no se distribuye de una manera uniforme.

Hay regiones relativamente conservadas tales como las regiones 5' y 3' no traducidas (NCRs), donde se requieren secuencias específicas y determinadas estructuras secundarias de RNA para las funciones de replicación y de traducción. Las glicoproteínas E1 y E2 y la proteína no estructural NS5A muestran más variabilidad dentro del genoma, en particular, algunas regiones hipervariables que también muestran un rápido cambio en la secuencia de aminoácidos en el tiempo. Parte de esta variabilidad surge a través de la selección por eventos asociados con el escape inmunológico (Farci 2011; Simmonds 2004).

Para analizar la dinámica evolutiva de las cuasiespecies intrahospedero en diferentes regiones del genoma de HCV, para un mismo paciente crónico se construyeron 2 dos datasets de secuencias nucleotídicas obtenidas a través del tiempo (16 y 18.2 años), para cada paciente, tomadas de los trabajos de Ramachandran, uno de los dataset corresponde a la región NS5A, y el otro de la región HVR1/E2, perteneciente a HCV del genotipo 1a, (Ramachandran *et al.* 2011).

El HCV ha desarrollado mecanismos de evasión de las defensas del hospedero que dependen de la actividad de la NS5A, (Tan y Katze 2001; Reyes 2002; MacDonald y Harris 2004).

La NS5A es una proteína no estructural que participan en la replicación del RNA viral. Otras funciones que le han sido atribuidas a NS5A son: la inactivación de PKR (Gale *et al.* 1997), el bloqueo de las vías de apoptosis mediante el secuestro de p53, la modulación de los niveles intracelulares de calcio, (Gong *et al.* 2001; Majumder *et al.* 2001; Szabo 2006) y la inducción de la secreción de interleuquina 8 (Polyak *et al.* 2001).

Enomoto y colaboradores fueron los primeros en sugerir que la heterogeneidad genética de un dominio específico en NS5A, denominada región sensible al interferón (ISDR), estaba relacionada con la respuesta al tratamiento en pacientes japoneses infectados con HCV de genotipo 1b, (Enamoto *et al.* 1995, 1996). Aunque este tema sigue siendo controvertido, otros estudios muestran hallazgos similares (Puig- Basagoiti *et al.* 2001; Giménez-Barcons *et al.* 2001; Witherell y Beineke 2001). Estudios recientes basados en meta-análisis de datos de la secuencia de ISDR de pacientes infectados con HCV genotipo 1b y su respuesta virológica al tratamiento con interferón, demostraron que el efecto del ISDR sobre la respuesta está universalmente presente, pero parecía ser más fuerte en los pacientes japoneses. Resultados discrepantes entre los estudios con pacientes japoneses y no japoneses pueden explicarse por las diferencias en los regímenes de dosificación y un efecto diferencial dependiente de la dosis en las mutaciones del ISDR en la respuesta al tratamiento (Schinkel *et al.* 2004).

Además de las mutaciones en el ISDR, las mutaciones en el denominado dominio V3 de NS5A (Inchauspe *et al.* 1991), una región de función desconocida que parece estar bajo fuerte presión selectiva, relacionada con la respuesta a la terapia antiviral (Duverlie *et al.* 1998; Nousbaum *et al.* 2000; Murphy *et al.* 2002).

A su vez la región HVR1 juega un papel importante tanto en la entrada a la célula del HCV como en la evasión inmune (Guan *et al.* 2012). HVR1 contiene epítotos neutralizantes, y su modificación puede llevar a escape del virus a partir de anticuerpos neutralizantes pre-existentes (Van Doorn *et al.* 1995; Dowd *et al.* 2009). Se ha propuesto que las sustituciones en HVR1 son accionadas por la presión inmune, y desempeñan un papel importante en el establecimiento de la infección persistente y la progresión de la enfermedad (Liu *et al.* 2010). Tres microdominios diferentes se han propuesto recientemente en HVR1 (Guan *et al.* 2012). El primer microdominio

(residuos 14, 15 y 25-27) juegan un papel clave en la unión del HCV al receptor SR-B1 siendo los residuos en este microdominio indispensables para entrar del HCV en la célula (Guan *et al.* 2012).

Una vez alineadas las secuencias mediante MEGA 5,05 (Tamura *et al.* 2011), se utilizó el servidor web Datamonkey (Delport *et al.* 2010), para identificar el modelo evolutivo más adecuado que mejor se ajuste a nuestros datos. Para ello utilizamos la aproximación Bayesiana Monte Carlo con Cadenas de Markov (MCMC) implementada en el paquete BEAST2 (Stadler *et al.* 2013), como modelo de población se utilizó el modelo GTR+ γ con 20 millones de generaciones de MCMC, se testearon diferentes dinámicas poblacionales (tamaño poblacional constante, skyline bayesiano, crecimiento poblacional exponencial, expansional y logístico), y los resultados se examinaron con el programa TRACER (Drummond and Rambaut 2007). La incertidumbre estadística en los datos se refleja en la densidad de probabilidad mayor al 95 % (HPD). La convergencia se evaluó con los valores de ESS (tamaño efectivo de muestreo).

A los efectos de investigar como varía la sub-población de cuasiespecies a lo largo del tiempo, se analizó la diversidad media dentro de las sub-poblaciones en cada tiempo, mediante el MEGA 5,05 (Tamura *et al.* 2011).

5.1 Resultados del objetivo específico 2

5.1.1. Análisis Bayesianos de la NS5A y la HVR1/E2 de los paciente B y C.

Luego de realizar un análisis de 20 millones de generaciones de MCMC, usando el modelo GTR+ γ , un reloj molecular relajado (Drummond *et al.*, 2006), encontramos que el tanto la NS5A como la HVR1/E2 sigue un modelo de skyline bayesiano para ambas regiones analizadas.

Las tasas evolutivas fueron para la HVR1/E2 en el paciente B de 5.11×10^{-3} s/s/a y en el paciente A de 4.44×10^{-4} s/s/a. y para la NS5A fueron de 1.25×10^{-2} s/s/a para el paciente B y de $1,38 \times 10^{-3}$ s/s/a para el paciente C. Los skyline correspondientes se muestran en la Figura 27.

5.1.2 Análisis de diversidad media dentro de las sub-poblaciones a través del tiempo.

La diversidad de una cuasiespecie se refiere a la relación de los individuos dentro de la población. La medida básica de la diversidad es la distancia media de Hamming, es decir, el número de posiciones mutadas en una secuencia particular con respecto a una secuencia dominante, secuencia de consenso, u otra referencia. Más a menudo, la diversidad se mide usando algoritmos de la distancia genética tales como la métrica de Jukes-Cantor o Kimura-2-parámetros, que toma en cuenta factores adicionales, tales como las probabilidades de transición (sustituciones de purina a purina o pirimidina a pirimidina), transversiones (sustituciones de pirimidina a purina y viceversa). Mediante Compute Within Group Mean Distance en MEGA 5,05, utilizando la métrica de Kimura-2-parámetros, se analizó la diversidad media entre las sub-poblaciones, graficándose los valores obtenidos en cada muestreo. La gráfica muestra pues como varia la diversidad de la nube de cuasiespecies a lo largo del tiempo. (Figura 28).

Tanto para el paciente B como par el paciente C, la diversidad media entre las sub-poblaciones en la HVR1/E2 y de la NS5A muestran un patrón similar de variación a lo largo del tiempo (Figuras 28 y 29).

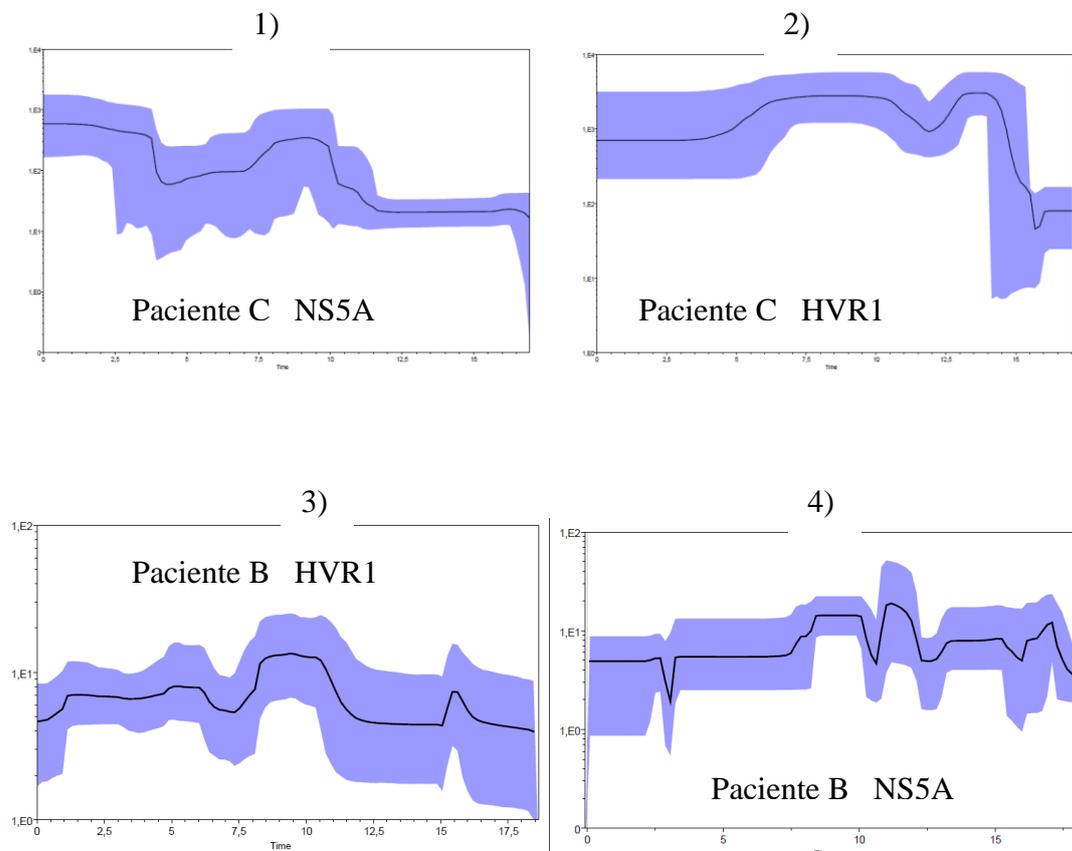


Figura 27. Skyline Bayesianos de la NS5A y la HVR1/E2 de los paciente B y C. Luego de realizar un análisis de 20 millones de generaciones de MCMC, usando el modelo GTR+ γ , con un reloj molecular relajado.

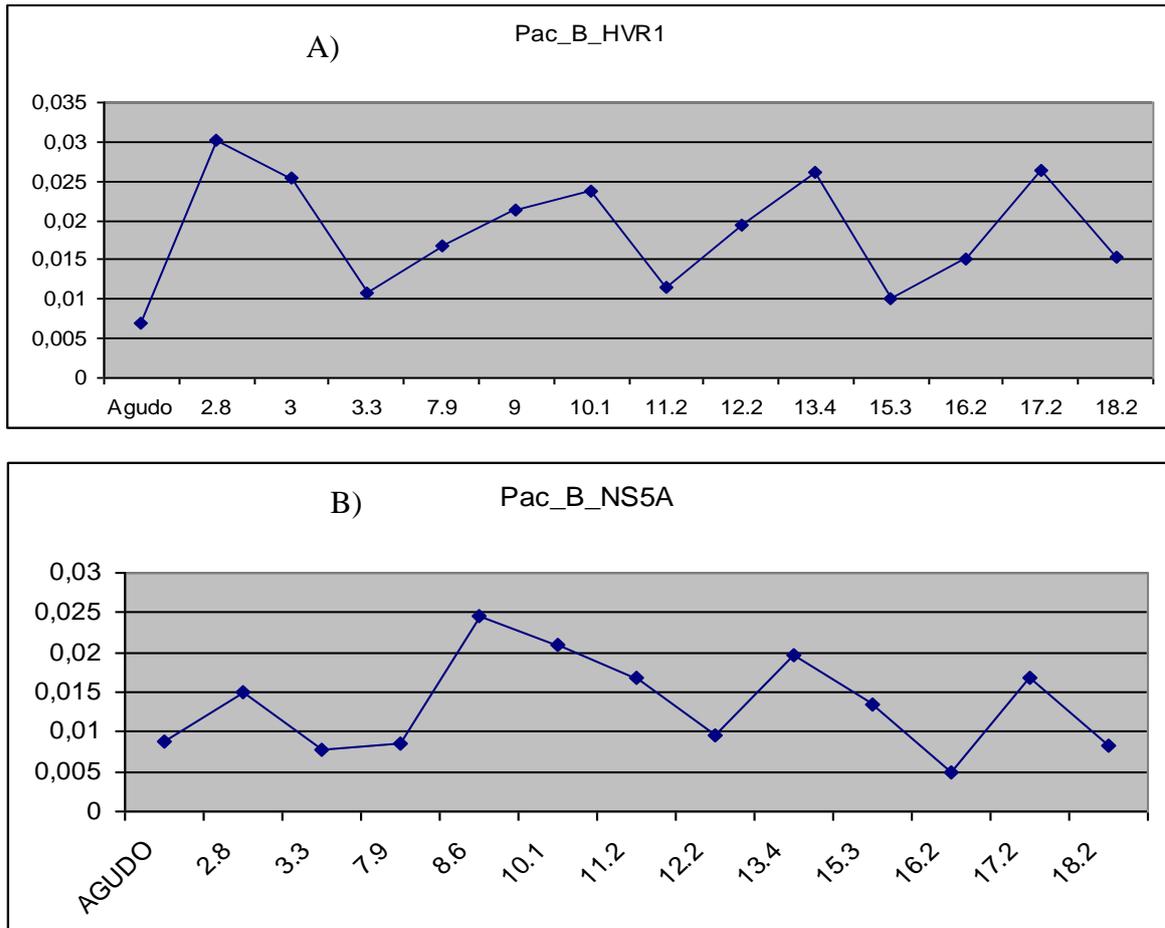


Figura 28. Gráfica de la diversidad media entre las sub-poblaciones de A) HVR1/E2 y B) NS5A, durante los diferentes períodos de tiempo estudiados, obtenida mediante Compute Within Group Mean Distance en MEGA 5,05 correspondiente al paciente B; utilizando la métrica de Kimura-2-parámetros. Se grafica la diversidad media de la sub-población en cada toma de tiempo en años desde la fase aguda.

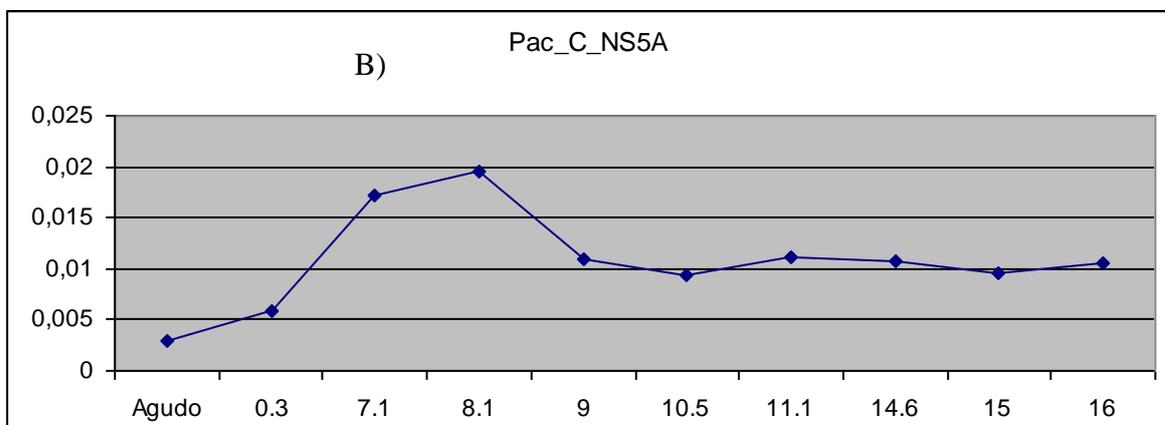
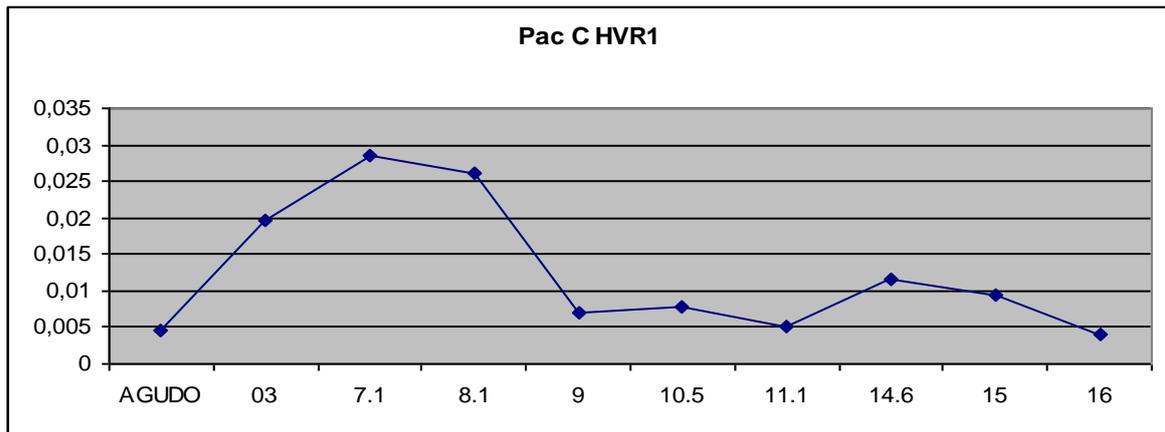


Figura 29. Gráfica de la diversidad media entre las sub-poblaciones de A) HVR1/E2 y B) NS5A, durante los diferentes períodos de tiempo estudiados, obtenida mediante Compute Within Group Mean Distance en MEGA 5,05 correspondiente al paciente C; utilizando la métrica de Kimura-2-parámetros. Se grafica la diversidad media de la sub-población en cada toma de tiempo en años desde la fase aguda.

5.2 Discusión del objetivo específico 2

5.2.1 Análisis de diversidad media dentro de las sub-poblaciones a través del tiempo.

Las gráficas de la diversidad media entre las sub-poblaciones de la HVR1/E2 y de la NS5A muestran un patrón similar dentro de cada paciente.

Si bien se trata de 2 regiones diferentes del genoma, la diversidad de la población de secuencias dentro de la nube de cuasiespecies experimenta variaciones similares, tal vez influenciada por variaciones en la presión inmune del hospedero a lo largo del tiempo.

6. Objetivo específico 3

Analizar la dinámica evolutiva de las cuasiespecies durante la transmisión materno-fetal.

La transmisión materno-fetal es la ruta predominante para la adquisición de la infección por HCV en niños en los países desarrollados; con una tasa de transmisión de aproximadamente el 5%, de esos niños casi el 80% desarrolla una infección crónica.

La persistencia de la infección en bebés y niños se asocia a un mínimo o escaso daño hepático, en los niños con infección crónica la población inicialmente es altamente homogénea, y la diversificación se observa durante el transcurso del tiempo, la diversificación de cuasiespecies del HCV parece ser un evento común en el desarrollo de hepatitis C crónica en la infancia.

Estudios recientes han mostrado evidencia que la lesión hepática se asocia con una población viral mono u oligoclonal, mientras que el daño hepático leve o escaso se correlaciona con la aparición temprana de cuasiespecies virales heterogéneas.

Por lo tanto, la evolución de las cuasiespecies que experimente el HCV en los niños infectados por vía perinatal puede correlacionarse con el daño hepático (Farci *et al*, 2006). Cuando la respuesta inmunitaria adaptativa específica emerge después de la infección primaria, el virus se enfrenta a nuevas restricciones ambientales que darán forma a su evolución a lo largo del curso de la enfermedad.

El poder conocer la forma en que evolucionan tempranamente las cuasiespecies podría ser útil para predecir la evolución de la enfermedad hacia la fase aguda. Como virus de RNA, el HCV posee una RNA polimerasa RNA dependiente que carecen de los mecanismos de corrección, haciendo la replicación del HCV altamente propensa a

errores (Moradpour *et al.* 2007). Como resultado, la población de HCV en cada paciente consta de genomas estrechamente relacionados pero no idénticos, formados por mutantes y recombinantes, sometidos a un continuo proceso de selección y competencia, lo que se conoce como cuasiespecies virales (Martell 1992; Pawlotsky 2006; Domingo *et al.* 2006).

La naturaleza dinámica de las cuasiespecies del HCV permite que el virus se adapte continuamente al desafío inmunológico del hospedero proporcionando un gran número de mutaciones potencialmente beneficiosas para permitir el escape inmunológico

La mayoría de los anticuerpos neutralizantes (nAbs) generados durante la fase aguda de la infección, tienen como blanco epítopes dentro de las proteínas virales, muchos de los cuales, han sido mapeados en las glicoproteínas de la envoltura E1 y E2, (Johansson *et al.* 2007; Kato *et al.* 1993; Keck *et al.* 2008; Meunier *et al.* 2008; Owsianka *et al.* 2005; Perotti *et al.* 2008; Shimizu *et al.* 1996). Estudios con chimpancés han demostrado que los nAb pueden proteger contra la infección por el HCV o reducir la severidad de la enfermedad (Verstrepen *et al.* 2011).

Tomando como referencia la cepa 1a H77 (AF011751), se ha identificado la región HVR1 de E2 (aa 384 y 410), como un importante blanco para nAb, ya que posee múltiple epítopes lineales, jugando así un importante papel en el reconocimiento de anticuerpos y en la evolución de la enfermedad, siendo necesaria para la unión al receptor scavenger clase B tipo I (SR-BI), una molécula de lipoproteína receptora involucrada en la entrada del HCV, (Bartosch *et al.* 2003.; Scarselli *et al.* 2002).

La HVR1 es necesaria para la interacción con el SR-BI, pues facilita la entrada viral a la célula, (Bartosch *et al.* 2005; Voisset *et al.* 2005). Sin embargo, la función de HVR1 en la infectividad puede variar entre los diferentes genotipos (Prentoe *et al.* 2011).

Un importante número de anticuerpos monoclonales (mAbs) específicos para HVR1 se han generado y caracterizado, la mayoría de los cuales reconocen epítopes en la región C-terminal de HVR1 (aa 396-407). En contraste, los mAb no neutralizantes se hallan asociados a la porción N-terminal de HVR1 (aa 384-395) (Hsu *et al.* 2003).

Identificándose así dos regiones inmunogénicas dentro de HVR1, donde la región C-terminal contiene los epítopes asociados a la neutralización, y la N-terminal los asociados a anticuerpos no neutralizantes.

Se ha propuesto que las sustituciones en HVR1 son accionadas por la presión inmune, y desempeñan un papel importante en el establecimiento de la infección persistente y la progresión de la enfermedad (Liu *et al.* 2010).

La HVR1 del HCV contiene 3 microdominios funcionales independientes, (Guan *et al.* 2012). Para la cepa H77 del genotipo 1a, el primer de microdominio incluye cinco residuos en las posiciones 14, 15, y 25-27 y juega un papel clave en la unión de la proteína de la envuelta de HCV al receptor SR-BI, juegan un papel clave en la unión del HCV al receptor SR-B1 siendo los residuos en este microdominio indispensables para entrar del HCV en la célula (Guan *et al.* 2012).

Los nueve residuos a través de las posiciones 16-24 constituyen el segundo microdominio, que contiene el epítipo de neutralización y parece ser necesario para mejorar la infectividad del HCV y le confiere resistencia a anticuerpos neutralizantes dirigidos a dos epítopos fuera de la HVR1, también juega un papel importante en la unión de HCV al heparán. El tercer microdominio de la HVR1 incluye los aminoácidos 1-13, es prescindible para la entrada en la célula por HCV, y afecta a la infectividad del HCV mediante la modulación de la unión de la proteína de envoltura a SR-BI. Las mutaciones en este microdominio pueden conferir resistencia a los anticuerpos de HCV en HVR1.

Esto muestra la importancia de la HVR1 en la mediación de la entrada del HCV a la célula, en la evasión inmune, y en la neutralización mediada por anticuerpos. Estudios recientes han demostrado que HVR1 es capaz de enmascarar epítopos nAb en E2, dado que mutantes con delección de HVR1 son mucho más susceptibles a la neutralización por un panel de mAbs humanos y sueros dirigidos contra el sitio de unión de la CD81-E2 (Bankwitz *et al.* 2010; Prentoe *et al.* 2011). Esto probablemente es debido al enmascaramiento de los sitio de unión con CD81. Por lo tanto, HVR1 puede funcionar para proteger la neutralización de los determinantes de la entrada viral dentro de E2 (Bankwitz *et al.* 2010). Esta naturaleza un tanto contradictoria, ha llevado a sugerir que HVR1 funciona como un señuelo inmunológico, estimulando una fuerte respuesta de anticuerpos hacia HVR1 que no da lugar a la eliminación del virus, pero que impulsa la selección de mutantes de escape (Ray *et al.* 1999).

Por lo cual mutaciones dentro de ésta región constituirían un importante mecanismo para la evasión a la respuesta inmunogénica del hospedero.

Muchos virus envueltos, incluyendo el virus herpes simple 1, y el sarampión, utilizan transmisión directa célula a célula, en un intento de evadir la respuesta inmune (Mothes *et al.* 2010).

Recientemente, se ha encontrado que el HCV es capaz de tener transmisión directa célula a célula, lo que lo hace en gran parte resistente a la neutralización por anticuerpos (Timpe *et al.* 2008; Witteveldt *et al.* 2009), aunque los mAb de rata 9/27 (dirigida contra los aa 396-407 en HVR1) y 11/20c (dirigido residuos de aa 412-423 y 436-447 en la unión a CD81) son capaces de inhibir parcialmente la transmisión célula a célula (Brimacombe *et al.* 2011). Por lo cual mutaciones en esta región podrían facilitar la transmisión célula a célula del virus.

Para analizar la dinámica evolutiva de las cuasiespecies intra-hospedero durante la transmisión materno-fetal, se construyó un dataset con secuencias aminoacídicas presentes en un paciente infectado con HCV genotipo 4d, durante una transmisión materno-fetal con un seguimiento de 93 meses, tomadas de los trabajos de Farci, las secuencias analizadas pertenecen parcialmente a las proteínas E1/E2 y comprenden las regiones HVR1 y HVR2 abarcando desde el aminoácido 323 a 498, tomando como referencia a la cepa H77c (AF011751), (Farci *et al.* 2006).

Una vez alineadas las secuencias aminoacídicas mediante MEGA 5,05 (Tamura *et al.* 2011), se compararon las secuencias encontradas en el momento de la transmisión con las encontradas durante los 12 muestreos a lo largo de los 93 meses del estudio, se utilizó el servidor web Datamonkey (Delport *et al.* 2010), para identificar el modelo evolutivo que mejor se ajuste a nuestros datos y se construyeron los árboles filogenéticos correspondientes, además se realizó un estudio de redes de los haplotipos, para lo cual se construyó la red de la población de cuasiespecies utilizando los enfoques de reducción de la mediana de unión con Network 4.6 (disponible en: <http://www.fluxus-engineering.com>). (Vease Materiales y Métodos)

También se compararon las secuencias consenso presentes durante todo el período de estudio, y con el fin de identificar epítopes a lo largo de las secuencias analizadas se utilizó el programa BepiPred, del servidor Datamonkey (Delport *et al.* 2010). Se mapearon las mutaciones encontradas con la finalidad de asociarlas a los epítopes predichos.

Se mapearon las sustituciones aminoacídicas encontradas con la estructura correspondiente, tomada de la Protein Database, a fin de identificar su ubicación mediante el programa PDB ProteinWorkshop 3.6 (Moreland *et al.*, 2005).

6.1 Resultados del objetivo específico 3

6.1.1 Análisis filogenéticos y de redes de los haplotipos durante las transmisión materno-fetal

Un análisis detallado de las secuencias aminoacídicas presentes en el paciente durante la transmisión materno-fetal de HCV genotipo 4d, con un seguimiento de 93 meses (Farci *et al.* 2006), revela la presencia de mutaciones dentro de la proteína de envoltura E2 asociadas a epítopes de nAbs, las que se fijan inmediatamente después de la seroconversión, las que brindan a la nueva secuencia la capacidad de evadir la acción de los nAbs, brindándoles una ventaja evolutiva a través del tiempo.

En el momento de la transmisión materno-fetal, 10 haplotipos fueron identificados, 3 de los cuales se hallaban con frecuencia superior a 1; el árbol filogenético reveló la existencia de 2 clados, perteneciendo 2 de las secuencias repetidas a uno de esos clados y la otra secuencia al otro clado; a las que denominamos secuencias fundadoras 1 y 2 respectivamente. Estas fueron: 813_F1, 8 veces; 834_F1, 7 veces; y 835_F2, 12 veces (Figura 30 a).

La construcción del árbol filogenético con los 49 haplotipos encontrados a través de los 12 muestreos abarcando los 93 meses, revela a las secuencias llamadas fundadoras 1 en un grupo alejado, solo emparentadas con las secuencias presentes en la primera toma, mientras que la secuencia llamada fundadora 2 se halla asociada a todas las demás secuencias presentes a partir del mes 7 y durante los 93 meses, (Figura 30 b). La red nucleotídica de los haplotipos a lo largo de los 93 meses también muestra dichas relaciones (Figura 31).

La segunda toma, realizada a los 3 meses, de un total de 31 muestra solo se presentan 2 haplotipos, uno de ellos asociado a la secuencia 813_F1 en 30 oportunidades.

A los 5 meses se produce la seroconversión

La tercera toma realizada a los 7 meses de un total de 40 muestras, se presentaron 4 haplotipos, uno de los cuales corresponde a la secuencia 835_F2 en 37 oportunidades.

Es decir en el momento de la transmisión 3 haplotipos se encontraban en forma mayoritaria, a los 3 meses uno de ellos, asociado a 813_F1 predomina, en 30 de las 31

secuencias presentes, no hallándose la secuencia 835_F2, pero luego de la seroconversión (mes 5), en la tercera toma correspondiente al mes 7 vuelve a hallarse el haplotipo asociado a 835_F2 en 37 de las 40 secuencias y a partir de allí permanece siendo predominante hasta el mes 93 (Figura 32).

6.1.2 Predicción de epítopes

Utilizando el programa BepiPred, para dichas secuencias se identifican 2 regiones de 6 aa de longitud cada una, asociadas a posibles epítopes dentro de HVR1 y HVR2, en ambas regiones se identifican 3 cambios aminoacídicos entre los haplotipos correspondientes a las Fundadoras 1 y 2 (Figura 33).

Las mutaciones de la región HVR1 pertenecen a los 3 microdominios de la misma. Dentro de dicha región se identifican 7 cambios de aminoácidos; el aminoácido 408 correspondiente al microdominio 1; los aa 404 y 405 dentro del microdominio 2; y los aa 384, 386, 387 y 391 dentro del microdominio 3; Cuatro de dichos cambios en la región C-terminal se asocian a epítopes de anticuerpos neutralizantes, y 3 en la región N-terminal reconocen epítopes de anticuerpos no neutralizantes, (Figura 34).

En el motivo altamente conservado entre todos los genotipos y asociado a epítopes reconocidos por anticuerpos neutralizantes, denominado EP II

430-NESLNTGWLAGLFYQHK-446

se identifica un cambio en el aa 440, el cual corresponde al centro de la α -hélice que reconoce a los nAb, (Figura 35).

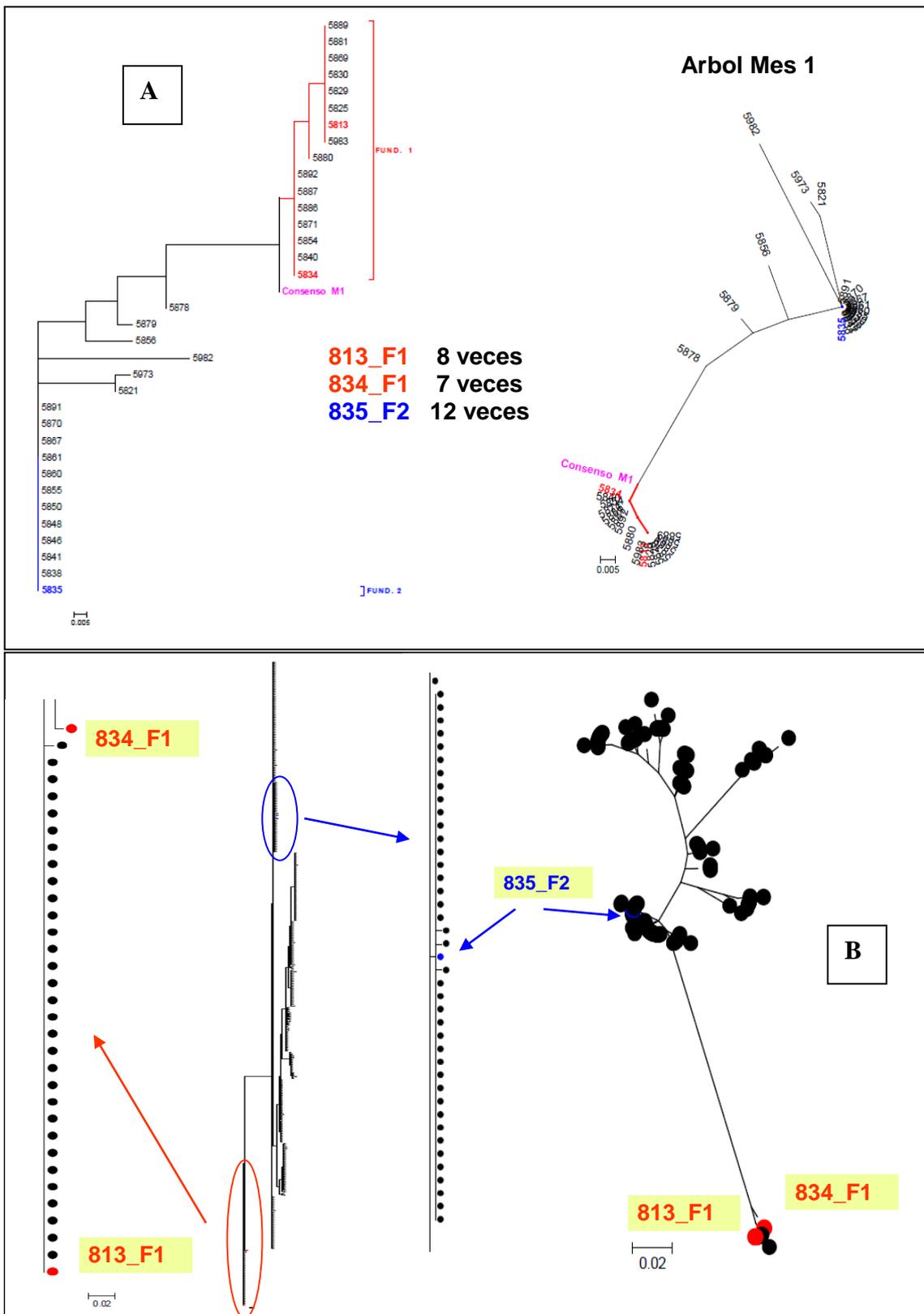


Figura 30. A) Árbol filogenético correspondiente a los haplotipos encontrados en el primer mes, se identifican 2 clados, las secuencias identificadas como 813_F1 y 834_F1 (en rojo), pertenecen a 1 de los clados y la 835_F2 pertenece al otro. En violeta se marca la secuencia consenso correspondiente al mes 1. B) Árbol filogenético de los haplotipos encontrados durante los 93 meses.

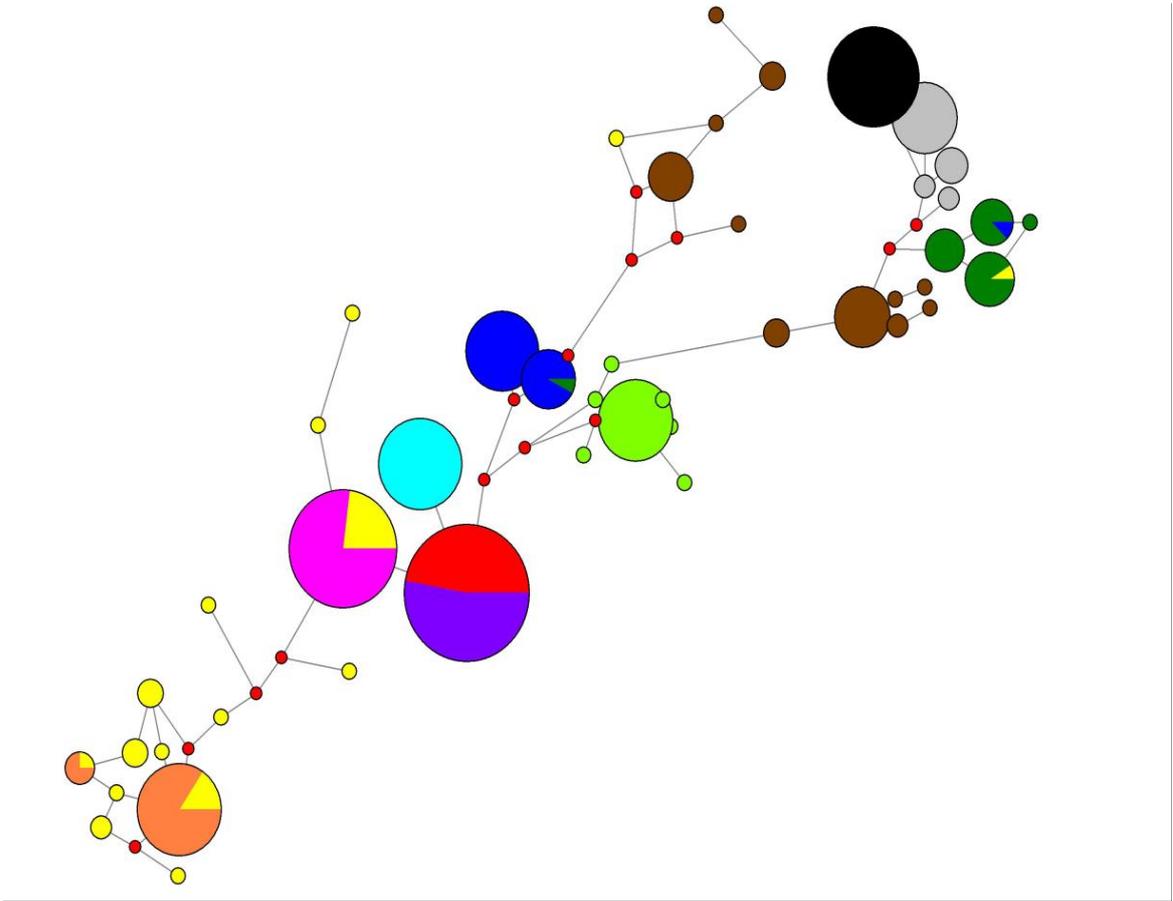


Figura 31. Red correspondiente a los 92 haplotipos encontrados durante los 93 meses del análisis. Mes1 amarillo; mes 3 ocre; mes 7 lila; mes 9 rojo; mes 15 violeta; mes 19 celeste; mes 41 azul; mes 46 verde claro; mes 55 verde oscuro; mes 67 marrón; mes 73 gris y mes 93 negro. El tamaño representa el número de secuencias presentes de cada haplotipos.

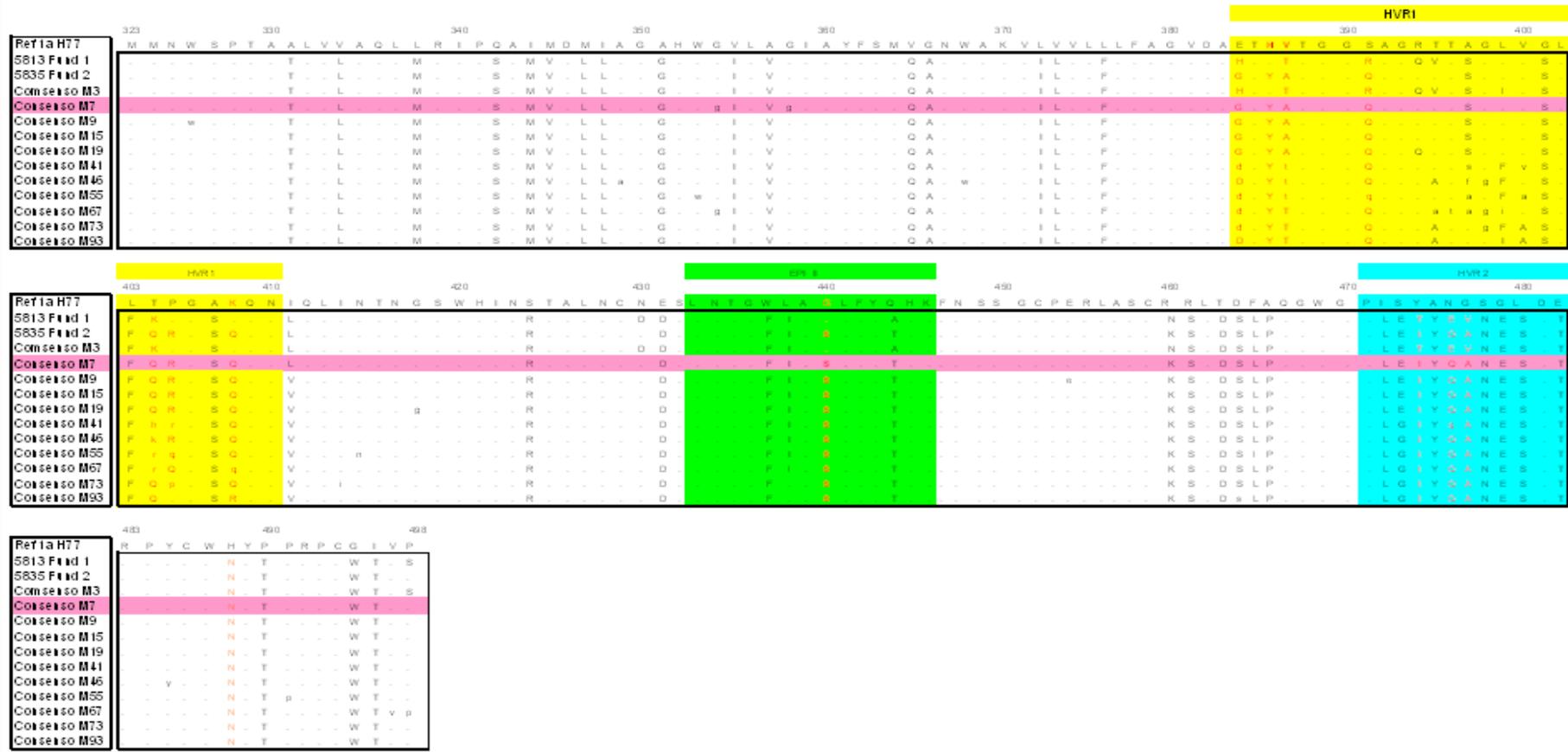


Figura 32.- Muestra el alineamiento de las secuencias consenso circulantes en cada unidad de muestreo, como se puede apreciar la mayoría de los cambios aminoacídicos se producen entre el mes 3 y el mes 7, período en el que se produce la seroconversión, se puede apreciar que muchos de esos cambios constituyen una reversión hacia la secuencia F_1, mientras que la secuencia consenso del mes 3 es mas similar a F_2.

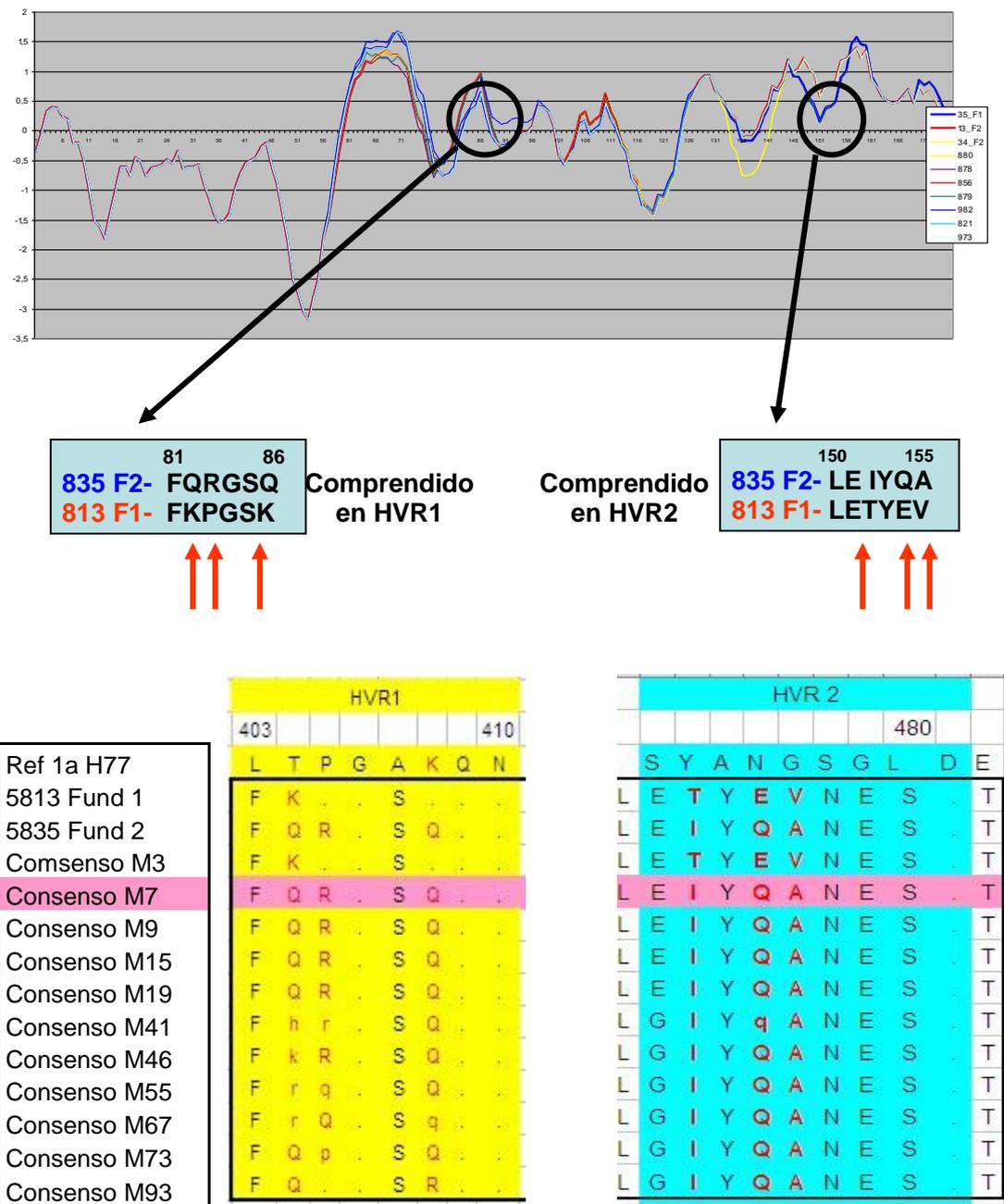


Figura 33. Arriba predicción de epitopes mediante el programa BepiPred se identifican 2 regiones de 6 aa de longitud cada una, asociadas a posibles epítopes dentro de HVR1 y HVR2, en ambas regiones se identifican 3 cambios aminoacídicos (flechas rojas) entre lo haplotipos correspondientes a las secuencias denominadas Fundadoras 1 y 2. Los números indican las posiciones de los aminoácidos dentro de la secuencia estudiada. 81-86 corresponden a las posiciones 403-409 dentro de la poliproteína; 150-155 corresponde a 472-477 de la poliproteína. Abajo Secuencias consenso de cada una de las tomas y las secuencias fundadores 1 y 2 correspondientes a dichas regiones.

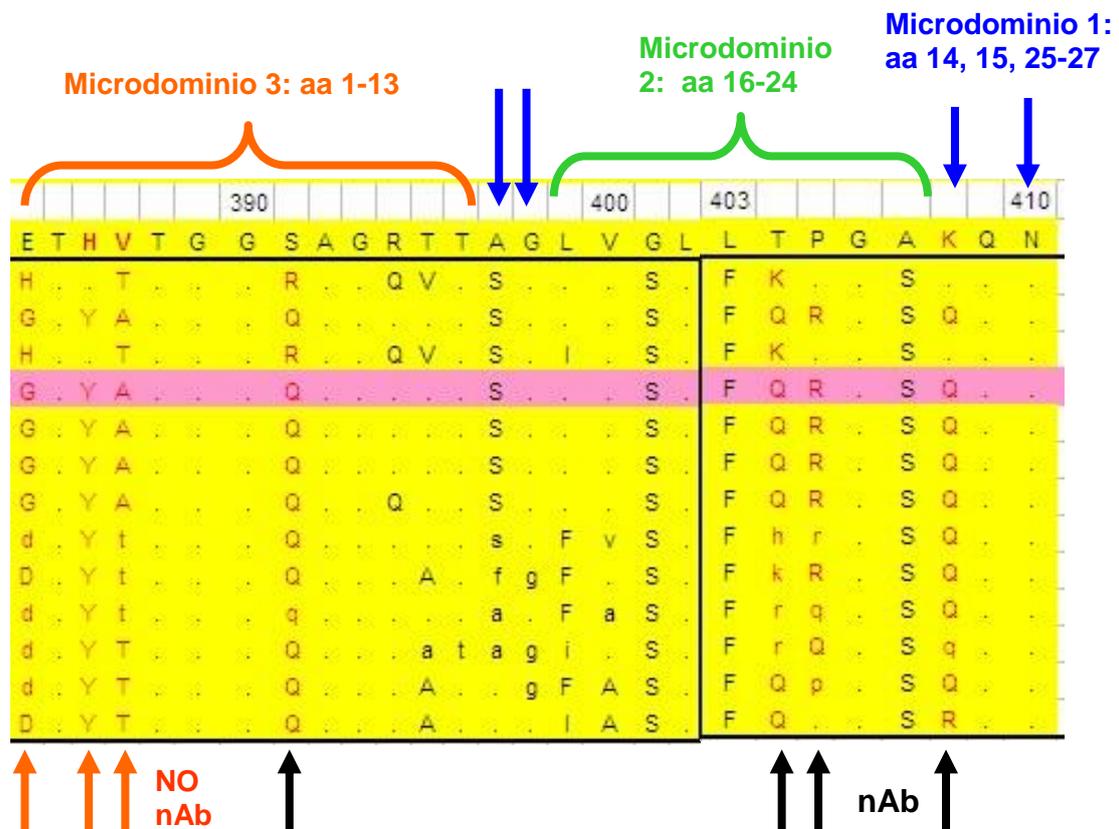


Figura 34. Se identifican los 7 cambios de aminoácidos entre las secuencias fundadoras 1 y 2 y su presencia en las secuencias consenso a lo largo del tiempo; el aminoácido 408 correspondiente al microdominio 1; los aa 404 y 405 dentro del microdominio 2; y los aa 384, 386, 387 y 391 dentro del microdominio 3; Los cambios en la región C-terminal se asocian a epítopes de anticuerpos neutralizantes (flechas negras), y los de la región N-terminal reconoce epítopes de anticuerpos no neutralizantes (flechas rojas).

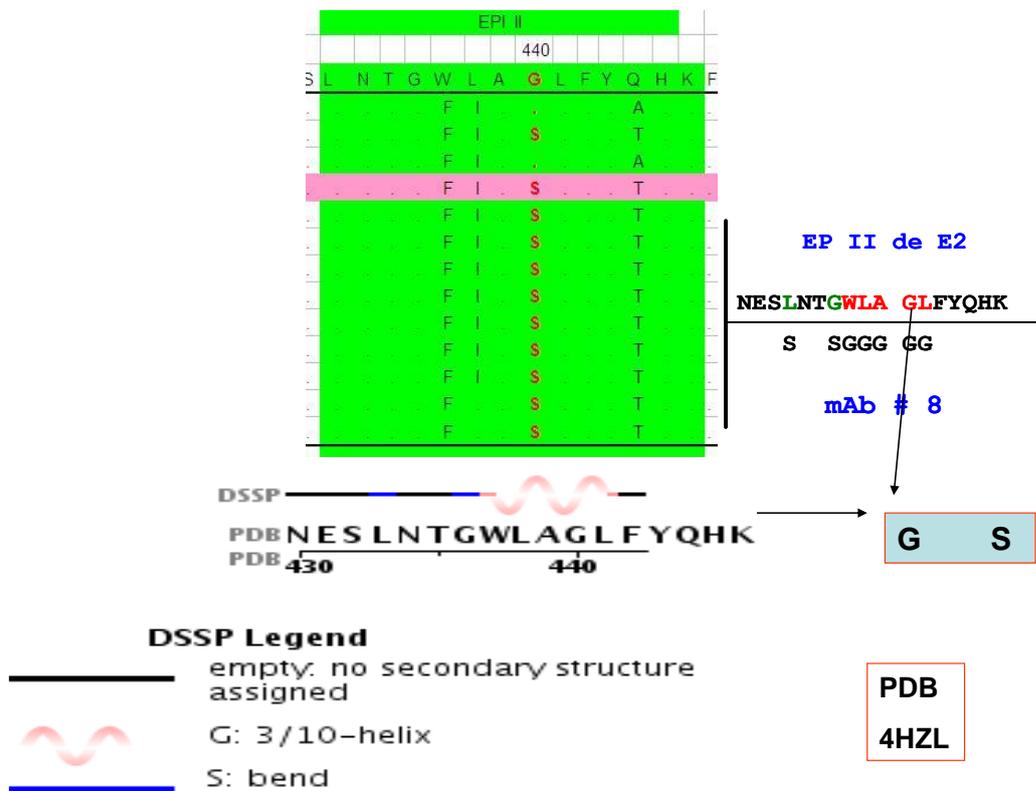


Figura 35. El mapeo sobre la estructura de la E2 de HCV en la Protein Data Base 4HZL Muestra la ubicación del aminoácido 440 sobre la α -hélice de la región conocida como EP II, que reconoce a los nAb. El cambio G-S entre las secuencias fundadoras 1 y 2.

6.2 Discusión del objetivo específico 3

6.2.1 Análisis filogenético y redes de los haplotipos durante las transmisión materno-fetal

Los análisis filogenéticos y de redes de las secuencias encontradas en el momento de la transmisión materno-fetal (1 mes del nacimiento), revelaron la presencia de 10 haplotipos diferentes, constituyendo claramente 2 clusters separados, identificándose una secuencia mayoritaria en cada uno de esos clusters (llamados fundadoras 1 y 2), en la siguiente toma realizada a los 3 meses del nacimiento, sólo 2 haplotipos fueron identificados, asociados al cluster de la fundadora 1, uno de ellos es encontrado en 30 de las 31 secuencias presentes, sin embargo en la tercera toma (a los 7 meses del nacimiento) y habiéndose producido la seroconversión (mes cinco), se identifican 4 haplotipos, asociado al cluster de la fundadora 2, uno de ellos es encontrado en 37 de las 40 secuencias presentes.

A partir de esa toma y hasta el final del estudio, (93 meses del nacimiento), todas las secuencias encontradas se asocian al cluster de la fundadora 2.

Esto hace suponer que luego de la transmisión materno-fetal, el pool cepas asociadas a la fundadora 1 encuentra condiciones que le son mas favorables y abarcan todo el espacio, sin embargo las cepas asociadas al otro cluster (fundadora 2), deben de seguir replicándose en bajas tasas, y una vez producida la seroconversión, lo que conlleva a un cambio en las presiones inmunes, estas se encuentran mejor adaptadas al nuevo ambiente y vuelven a ocupar el espacio, desplazando a las del cluster de la fundadora 1.

Una vez más se observa la importancia de los genomas memorias en la transmisión de la infección.

6.2.2 Predicción de epítopes

Mediante el programa BepiPred se identifican 2 regiones de 6 aa de longitud cada una, asociadas a posibles epítopes dentro de HVR1 y HVR2, ambas regiones se asocian a cambios aminoacídicos entre los haplotipos correspondientes a las secuencias denominadas Fundadoras 1 y 2. Mostrando la importancia evolutiva de estos cambios en dichas regiones.

6.2.3 Análisis de las mutaciones asociadas a los epítopes que reconocen nAbs.

El reconocimiento de la existencia de mutaciones en posiciones asociadas con epítopes que reconocen nAbs y no nAbs serían las responsables capaces de conferir al haplotipo 835_F2 una importante ventaja evolutiva sobre el haplotipo 813_F1 una vez producida la seroconversión.

Los cambios aminoacídicos encontrados en la HVR1 están asociados al los 3 microdominios de la HVR1, afectando epítopes asociados a nAbs.

6.2.4 Mapeo de la sustitución en el aa 440

El cambio en el aminoácido 440 se mapeo sobre la estructura de la HVR1 en la Protein Data Base (4HZL), constatando que el cambio G-S entre las secuencias fundadoras 1 y 2, se ubica sobre la α -hélice de la región altamente conservada entre todos los genotipos conocida como EP II, que reconoce a los nAb. Pudiendo pues este cambio resultar beneficioso a fin de poder escapar a las defensas inmunitarias del hospedero.

En resume los cambios encontrados asociados a los nAbs, que conferirían esta ventaja evolutiva se aprecian en la Figura 36.

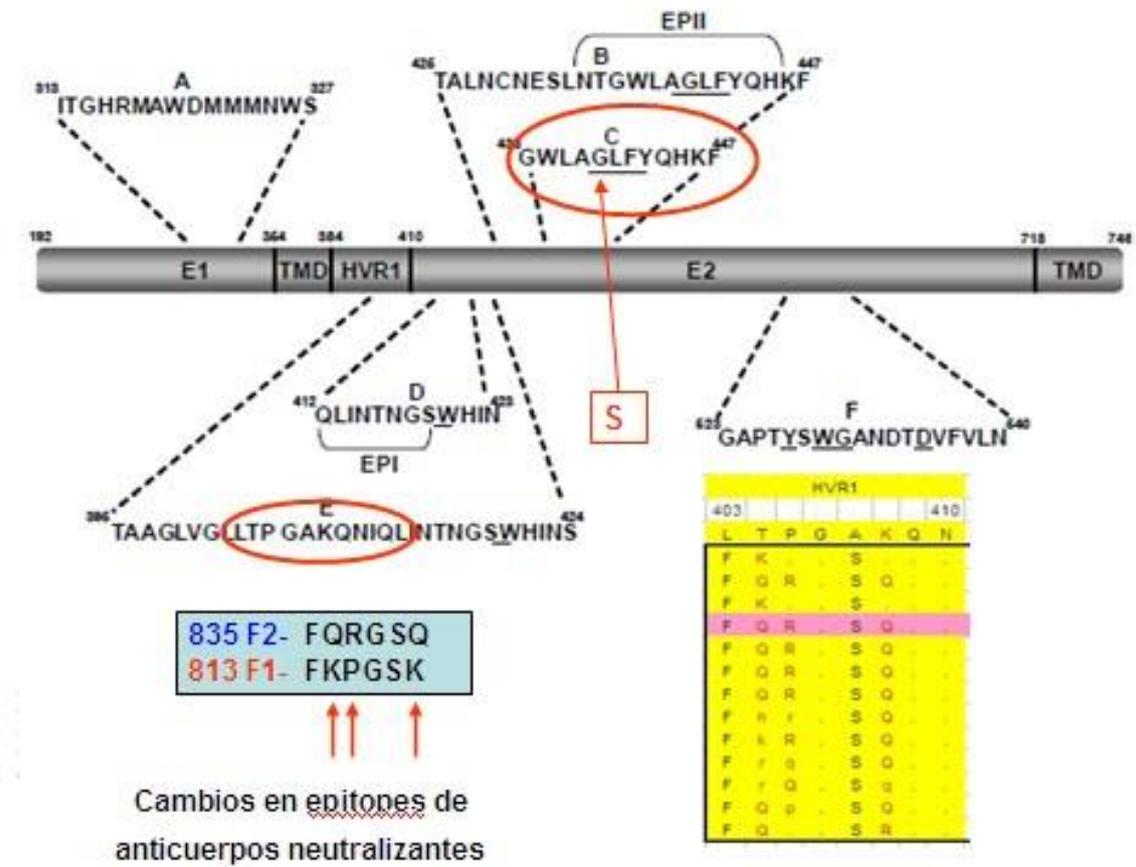


Figura 36. Resume los cambios encontrados asociados a los nAbs, que conferirían una ventaja evolutiva.

7. Objetivo específico 4

Analizar la dinámica evolutiva de las cuasiespecies por pirosecuenciación.

Recientemente se han desarrollado plataformas de secuenciación de nueva generación (NGS), que hacen posible investigar cuasiespecies virales en mucho mayor detalle.

La secuenciación masiva ha permitido a los investigadores secuenciar una fracción mucho mayor de las cuasiespecies. Este aumento en los genomas de la muestra se produce a expensas de una menor cobertura en la longitud de las secuencia (menos de 200 pares de bases), (Eriksson *et al.*, 2008).

Con el fin de profundizar en la dinámica evolutiva de las cuasiespecies intra-hospedero, hemos realizado un análisis coalescente bayesiano de la región HVR1 de secuencias de una población de cuasiespecies que circula en un paciente crónico infectado con HCV de genotipo 1b, obtenido recientemente por ultrasecuenciación profunda con ROCHE 454, (Kamila Caraballo-Cortés *et al.* 2013).

Con las plataformas de NGS, es posible investigar cuasiespecies virales en mucho mayor detalle. Su alto rendimiento permite la generación de millones de lecturas en una sola carrera de secuenciación, lo que facilita el análisis en profundidad (Kamila Caraballo-Cortés *et al.* 2013). Estos nuevos métodos pueden detectar las variantes presentes en bajas frecuencias, lo que no es detectada por los métodos de secuenciación estándar (Barzon *et al.* 2011). Sin embargo, con el fin de hacer fiable la reconstrucción de las cuasiespecies virales a partir de los datos ruidosos obtenidos por NGS, se requiere un análisis de datos adecuado (Beerenwinkel 2011).

Los análisis se realizaron a partir de 100 secuencias de aminoácidos de la HVR1 de HCV obtenidos por Carballo y col. Estas secuencias representan variantes del HCV cuyas frecuencias en la población osciló entre 10,49 % (el más abundante) y 0,10 % (las variantes menos abundantes). Estas secuencias también representan 100 diferentes haplotipos presentes en la población de cuasiespecies de HCV (llamados seq.1 a Seq.100) (Kamila Caraballo-Cortés *et al.* 2013).

Tres razones principales han sido tenidas en cuenta para la inclusión de estas secuencias en la construcción del conjunto de datos.

En primer lugar, esta región fue elegida dado que la proteína codificada se halla bajo presión de selección constante de las respuestas inmunes del hospedero (Di Lorenzo *et al.* 2011; Guglietta *et al.* 2005).

En segundo lugar, las secuencias de HVR1 se obtuvieron por secuenciación de ultra- profunda dieron lugar a 76.332 secuencias individuales (Kamila Caraballo-Cortés *et al.* 2013).

En tercer lugar, la corrección de errores en la reconstrucción de haplotipos para inferir las secuencias de HVR1 de las cuasiespecies fue hecho de forma probabilística usando un enfoque bayesiano, por medio de la utilización de el programa (Shorah) (Kamila Caraballo-Cortés *et al.* 2013; Zagordi *et al.* 2011).

Una vez alineadas las secuencias mediante MEGA 5,05 (Tamura *et al.* 2011), se utilizó el servidor web Datamonkey (Delpont *et al.* 2010), para identificar el modelo evolutivo más adecuado que mejor se ajuste a nuestros datos. Se utilizó un enfoque bayesiano MCMC para estudiar el modo de evolución de la población de cuasiespecies del HCV tal como se aplica en el paquete BEAST2 (Stadler *et al.* 2013), como modelo de población se utilizó el skyline de bayesiano de nacimiento-muerte.

El análisis del árbol filogenético fue realizado por dos enfoques diferentes. En primer lugar, se utilizó el programa DensiTree (Bouckaert 2010). En segundo lugar, árboles de máximo credibilidad fueron generados usando el programa Tree Annotator del paquete BEAST y el programa FIGTREE v1.2.2 (disponible en <http://tree.bio.ed.ac.uk>).

Construimos la red de la población de las cuasiespecies utilizando los enfoques de reducción de la mediana de la red mediante el programa Network 4.6 (disponible en: <http://www.fluxus-engineering.com>). (Vease Materiales y Métodos).

Con el fin de obtener una perspectiva de las interacciones evolutivas entre los sitios de la HVR1 empleamos un modelo gráfico bayesiano (BGM) (Poon *et al.* 2007), mediante la utilización del servidor Datamonkey (Delpont *et al.* 2010).

Para comprender mejor las posiciones de co-evolución de los sitios de aminoácidos en la región HVR1, se modeló in silico mediante la implementación del programa 3D-JIGSAW v3.0 (Offman *et al.* 2008).

Una vez obtenida la estructura se visualizó usando el programa J Mol v14.0.4 (Hanson 2010).

7.1 Resultados del objetivo específico 4

7.1.1 Análisis coalescente bayesiano de las secuencias de cuasiespecies de HVR1 de HCV.

Los resultados mostrados en la tabla de la figura 37 son el resultado del análisis de 90 millones de pasos de la MCMC, utilizando el modelo de JTT + invariantes, un reloj relajado (Drummond *et al.* 2006), y el Skyline del modelo nacimiento-muerte (Stadler *et al.* 2013).

Como se puede observar en la tabla de la figura 37, la convergencia es apoyada por valores altos de ESS. Cuando se utilizó el modelo de JTT, una tasa media fue de $4,80 \times 10^{-2}$ sustituciones de aminoácidos por sitio por año, para la evolución de la región HVR1 de esta población de cuasiespecies de HCV, revelando una tasa significativa y alta de evolución en esta región del genoma del HCV.

7.1.2 Análisis del árbol filogenético de las secuencias de cuasiespecies de la HVR1 de HCV.

DensiTree proporciona un método rápido para el análisis cualitativo de los conjuntos de árboles (Bouckaert 2010), las áreas donde muchos de los árboles están de acuerdo en longitud y topología de las ramas se presentan como áreas de intensa coloración, mientras que las áreas con poco acuerdo aparecen como redes menos densas.

Como se puede ver en la figura 37, la población de cuasiespecies de HCV aislada del paciente se puede asignar a tres grupos principales diferentes. El haplotipo más frecuentemente representado en la población de cuasiespecies (Sec.1, con una frecuencia de 10,49%) se asigna al grupo 1, el haplotipo menos frecuente (Sec.100, con una frecuencia de 0,10%) se asigna al grupo 2. Curiosamente, un diferente grupo filogenético (grupo 3) se observa claramente, formado exclusivamente con haplotipos cuya abundancia oscila de 7,41% a 0,10 %. Estos resultados también sugieren una co-circulación temporal de los tres diferentes grupos (Figura 37).

Con el fin de confirmar estos hallazgos, se realizó un análisis mediante la construcción de un árbol filogenético de máxima credibilidad. Los resultados de estos estudios se muestran en la Figura 38.

Una vez más, se observaron tres principales grupos filogenéticos. Las cepas asignadas al mismo grupo en la Figura 1 utilizando el enfoque DensiTree también se agrupan juntas en el árbol de máxima credibilidad (Figura 38). A su vez, el grupo 3 ha ido evolucionando en un linaje genético diferente de los grupos 1 y 2, a partir de antepasados que existían aproximadamente alrededor de 1,7 años antes del tiempo de aislamiento. Se puede también observar una temporal co-circulación de los diferentes linajes.

7.1.3 Análisis de redes de mediana de la población de cuasiespecies de la HVR1 de HCV.

Con el fin de estudiar el grado de variabilidad genética y la evolución de las cuasiespecies de HCV, es aconsejable utilizar diferentes y complementarios enfoques. Por estas razones y con el fin de confirmar los resultados obtenidos por el árbol filogenético, se construyó una red mediana de la unión de las cuasiespecies presentes en la población. Hemos utilizado una opción reducida red mediana-unión por en lugar de una red mediana completa para mejorar la claridad (Figura 39).

Como se puede ver en la figura, aunque una red es en forma de estrella la secuencia más abundante (seq.1) no se encuentra en el centro de la red. Vectores de significativa distancias media la separar los otros dos grupos genéticos. Esto habla de nuevo de un alto grado de variabilidad genética entre las cuasiespecies que se encuentran en la población.

La secuencia más abundante en las cuasiespecies no es el centro de una red, lo que sugiere que diferentes linajes genéticos están presentes en la población de cuasiespecies de acuerdo con los análisis de árbol filogenético.

7.1.4 Co-evolución de sitios de aminoácidos en la HVR1 de las cuasiespecies de HCV.

La región HVR1 juega un papel importante tanto en la entrada a la célula del HCV como en la evasión inmune (Guan *et al.* 2012). HVR1 contiene epítomos neutralizantes dominantes, y su modificación puede llevar a escape del virus a partir de anticuerpos neutralizantes pre-existentes (Van Doorn *et al.* 1995; Dowd *et al.* 2009). Se ha propuesto que las sustituciones en HVR1 son accionadas por la presión inmune, y

desempeñan un papel importante en el establecimiento de la infección persistente y la progresión de la enfermedad (Liu *et al.* 2010).

Tres microdominios diferentes se han propuesto recientemente en HVR1 (Guan *et al.* 2012). El primer microdominio (residuos 14, 15 y 25-27) juegan un papel clave en la unión del HCV al receptor SR-B1 siendo los residuos en este microdominio indispensables para entrar del HCV en la célula (Guan *et al.* 2012).

Con el fin de profundizar en los sitios de co-evolución y la interacción entre el HVR1, se realizó un análisis mediante el modelo gráfico bayesiano (Poon *et al.* 2007). Los resultados de estos estudios se muestran en la Figura 40.

Curiosamente, fuertes asociaciones en la red consenso ocurrieron entre los residuos 12 y 15, y los residuos 13 y 14 (probabilidad posterior de sustituciones correlacionadas de 1,00). Otra fuerte componente de asociaciones dentro de la red de asociaciones con el residuo 12, se da con el residuo 26 (probabilidad posterior de 0,81). Esto habla de la posible relación de residuos en las posiciones 12 y 13 de la HVR1 del HCV para permitirle al virus cambiar entre combinaciones de residuos permitiéndole escapar del sistema inmune mientras que conserva su estructura y función.

Una interacción estructural se produce entre los residuos que cooperan en la formación y estabilización de estructuras de proteínas secundarias o terciarias (Poon *et al.* 2007).

Desafortunadamente, la estructura de E2 completa es desconocida en la actualidad. Por esa razón, con el fin de obtener una perspectiva de las posibles interacciones entre los sitios de HVR1, modelamos *in silico* a la HVR1 de la cepa referencia H77 HCV (genotipo 1) y se mapearon los sitios que interactúan en este modelo. Los resultados de estos estudios se muestran en la Figura 41.

El modelo 3D *in silico* de HVR1 describe una cola (aminoácidos posiciones 1-8), una α -hélice (posiciones 9-18), una segunda cola (posiciones 19-22), un segundo α -hélice (posiciones 23-24) y una cola terminal (posiciones 25-27) (Figura 41). Curiosamente, los residuos que se consideren asociados en la red de sitios de co-evolución, al igual que las posiciones 12 y 15 o la posición 13 y 14, están situados en una región α -hélice de la HVR1. Esto sugiere que las interacciones pueden permitir que los residuos sean reemplazados por otras combinaciones mientras que se mantiene la estructura ordenada.

Group	Parameter	Value ^a	HPD ^b	ESS ^c
HCV HVR1	Log likelihood	-1160.54	-1181.57 to -1140.99	331.32
Sequences	Posterior	-1146.66	-1184.7 to -1110.57	1627.99
	Prior	13.882	18.54 to 46.74	1414.53
	Mean Rate ^d	4.80×10^{-2}	2.16×10^{-2} to 7.00×10^{-2}	228.45
	Root age (years)	1.74	0.59 to 4.3	100.02

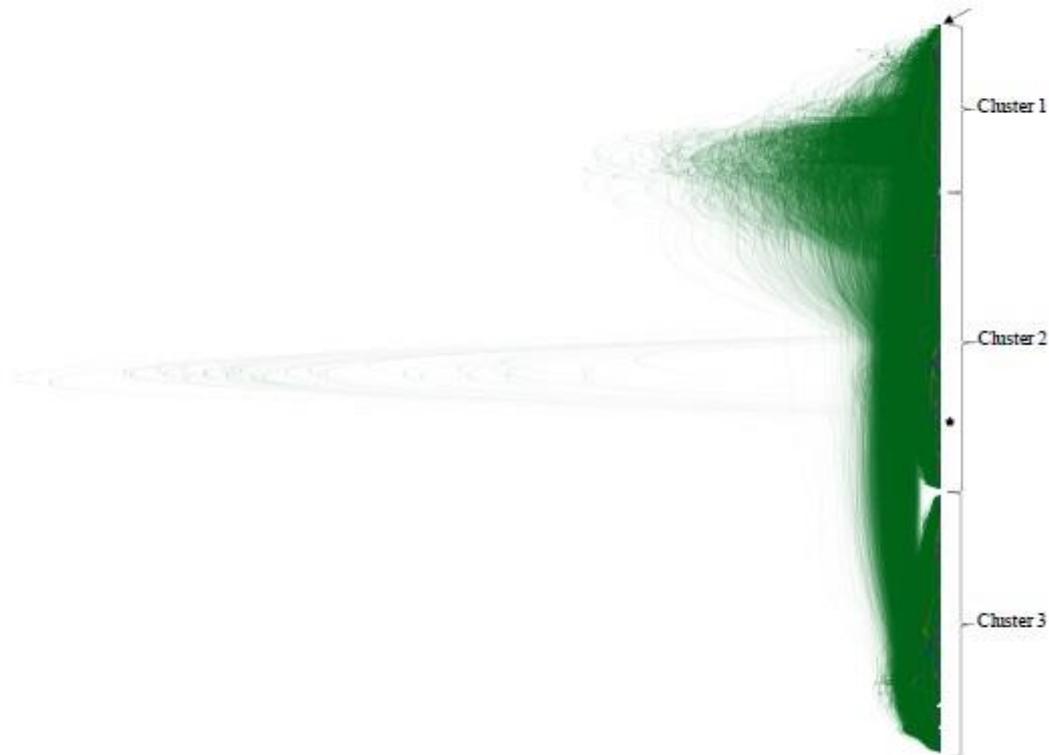


Figura 37. Análisis mediante DensiTree del árbol filogenético de HVR1 de genes E2 de las cuasiespecies del HCV. Se muestra el Consenso de más de 20.000 árboles. Las agrupaciones que representan diferentes sub-poblaciones de cuasiespecies están indicadas por los números en la parte derecha de la figura. Cluster 1 contiene 23 secuencias (Seq 1, 2, 13, 14, 17, 19, 23, 31, 39, 40, 44, 45, 55, 58, 59, 60, 61, 63, 64, 75, 77, 89 y 95); Cluster 2 contiene 40 secuencias (Seq 8, 9, 16, 26, 27, 29, 30, 32, 33, 34, 35, 36, 37, 41, 47, 48, 49, 52, 53, 54, 56, 62, 65, 68, 69, 71, 73, 74, 79, 81, 85, 86, 90, 93, 94, 96, 97, 98, 99, 100) y el Cluster 3 contiene 37 secuencias (Seq 3, 4, 5, 6, 7, 10, 11, 12, 15, 18, 20, 22, 24, 25, 28, 38, 42, 43, 46, 50, 51, 57, 66, 67, 70, 72, 75, 76, 78, 80, 82, 83, 84, 87, 88, 91, 92). Una topología dominante se observa para la mayor parte del árbol. Dentro de cada sub-población, se dan tres interpretaciones topológicas alternativas, azul para la más frecuente topología, rojo para la segunda, y verde claro para la tercera. La cepa más abundante (Sec. 1) se muestra por una flecha, y la menos abundante (Sec. 100) se muestra por una estrella.

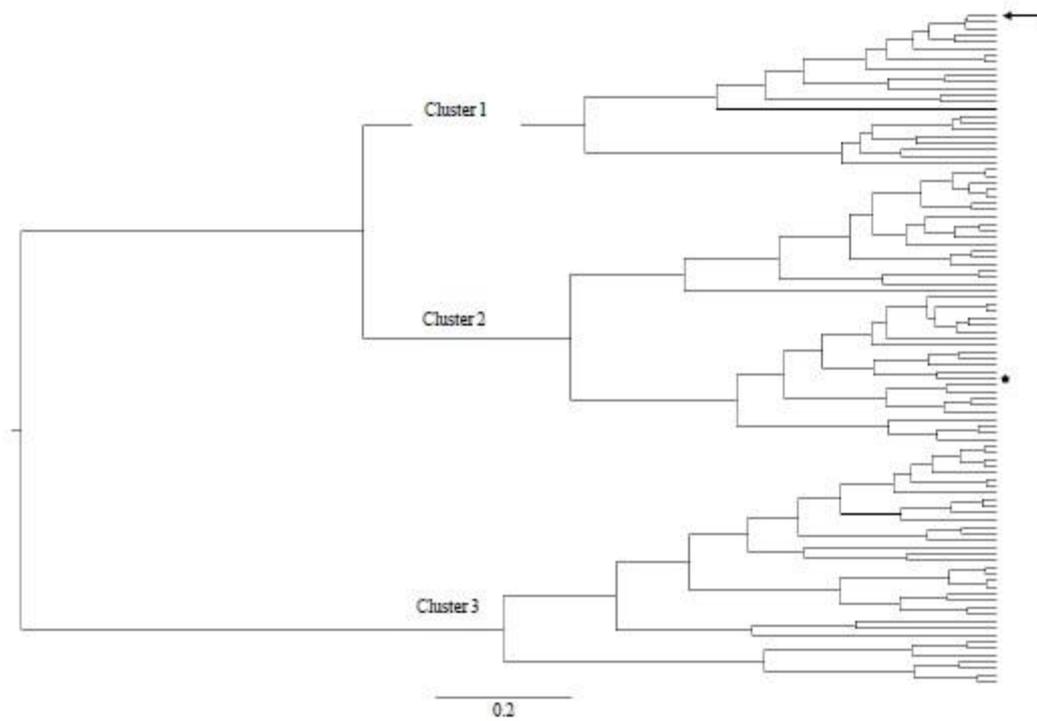


Figura 38 Análisis Bayesiano MCMC del árbol filogenético de HVR1 de genes E2 de cuasiespecies del HCV. Se muestra los cluster de máxima credibilidad obtenida mediante el modelo de JTT, el modelo de población horizonte de nacimiento-muerte y un reloj relajado (exponencial sin correlación). El árbol tiene en sus raíces un ancestro común más reciente (MRCA). La barra en la parte inferior muestra la edad del árbol en años. El resto se corresponde igual que la figura 37.

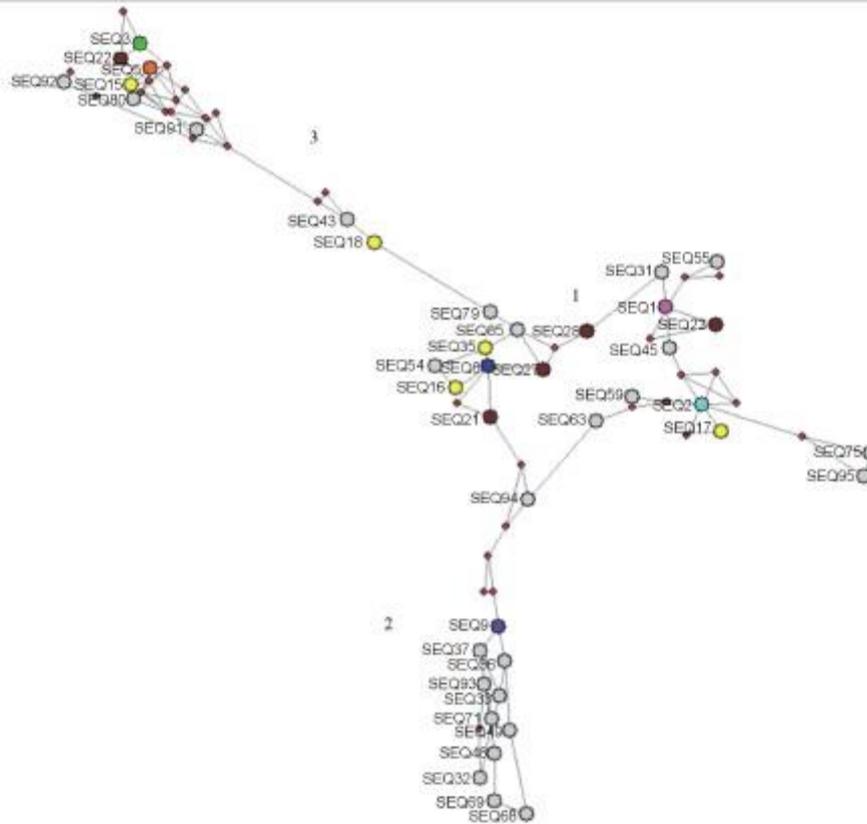


Figura 39. Análisis de redes de las cuasispecies intrahospedero de la HVR1-E2 del HCV. Se muestra una red de unión mediana reducida. En la red, cada nodo representa un haplotipo único dentro de la población de cuasispecies virales. La longitud del enlace representa las diferencias de aminoácidos entre los dos haplotipos diferentes. Las secuencias se muestran por su nombre al lado de los nodos. Los números representan las principales sub-poblaciones en la red. Los colores representan las frecuencias de abundancia relativa se encuentran en la población de cuasispecies. Seq.1 con 10,49 % (haplotipo más abundante) se muestra en fucsia, Sec.2, con 8,54%, se muestra en verde, Seq.5 con 4,29%, se muestra en color naranja; Sec.8 con 2,25%, se muestra en azul, la Seq.9 con 2,25%, se muestra en violeta; Seq.15-18 con 1,07%, se muestra en amarillo; Sec.21-23,26 y 27 con 0,80-0,65%, se muestra en marrón; haplotipos con menos del 0,62% de abundancia se muestran en gris.

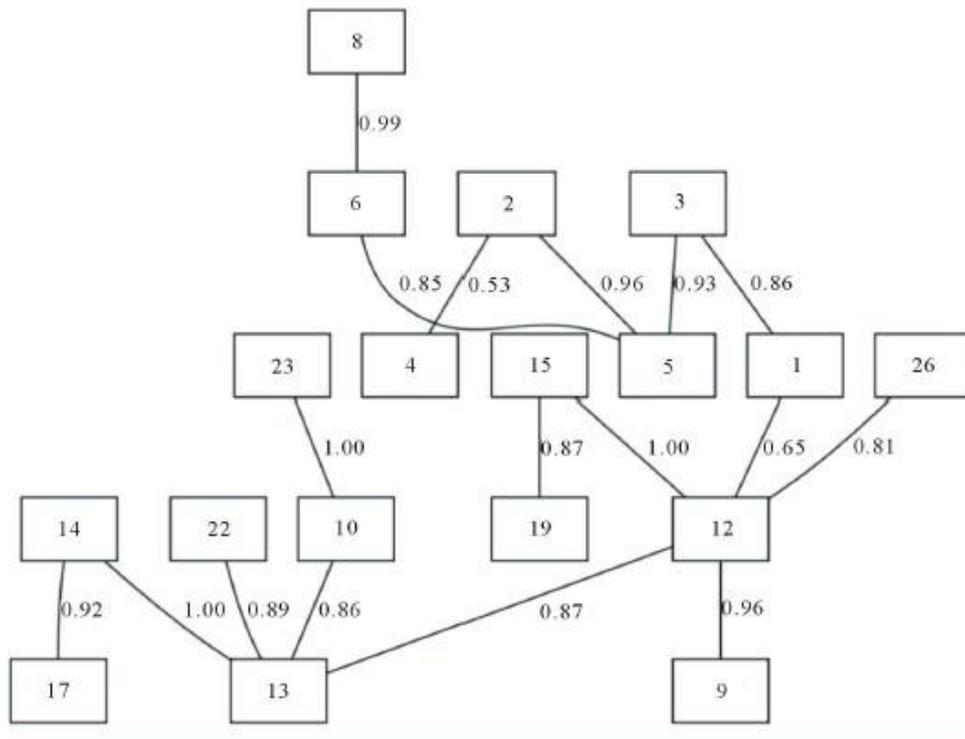


Figura 40. Red de Consenso de la co-evolución de los sitios de la región HVR1 de una población de cuasiespecies del HCV. Cada nodo, representado por un cuadrado, corresponde a un residuo en la HVR1, numerados de acuerdo con su posición relativa en la HVR1 del HCV aislado de H77 (genotipo 1). La probabilidad posterior de correlación de sustituciones entre cada par de sitios se indica mediante números entre los dos sitios.

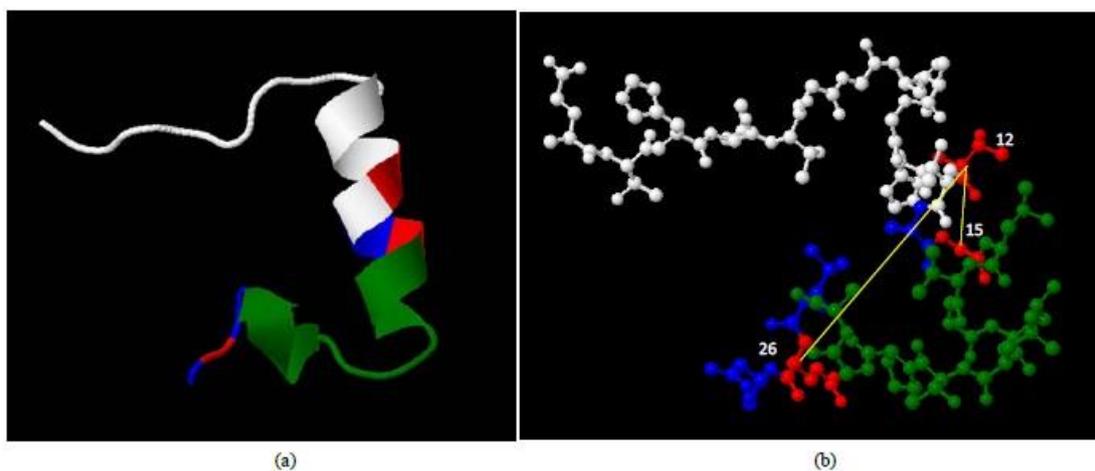


Figura 41. Modelo estructural 3D in silico de la HVR1 del HCV de la cepa referencia H77. Se muestra la estructura tridimensional de la molécula. Los microdominios 1 a 3 de la HVR1 se muestran en azul, verde y blanco, respectivamente. Se muestran los sitios de co-evolución de los aminoácidos 12 ,15 y 26 indicados en rojo y sus posiciones relativas en la estructura 3D se indica por las líneas amarillas. En (a) la estructura predicha de la HVR1 se muestra en un esquema de estilo de cintas; en (b) los aminoácidos en la molécula se muestran por palos y las bolas.

7.2 Discusión del objetivo específico 4

El HCV evoluciona a velocidades que están en el rango de 10^{-2} a 10^{-3} sustituciones por sitio por año (Lutchman *et al.* 2007). Siendo estas tasas evolutivas entre las más altas reportadas para los virus de RNA (Lutchman *et al.* 2007). El uso de métodos de secuenciación de profundidad se ha incrementado de manera significativa a través de los últimos años (Abe *et al.* 2013), y ha revelado una nueva perspectiva en los estudios evolutivos de los virus de RNA, lo que permite una capacidad sin precedentes para la caracterización de variantes de menor importancia dentro del espectro de las cuasiespecies virales (Jackowiak *et al.* 2014).

Este estudio nos ha permitido observar la evolución de la población de cuasiespecies en mayor detalle. La composición, la complejidad, y la amplitud del espectro de mutante de una población de cuasiespecies de HCV por secuenciación ultra-profunda nos muestra la capacidad de la población viral para encontrar las porciones del espacio de secuencia para mejorar su fitness (Domingo *et al.* 2012).

7.2.1 Análisis coalescente bayesiano de las secuencias de cuasiespecies de HVR1 de HCV.

Los análisis bayesianos proporcionan un método potente para la reconstrucción de las relaciones evolutivas (Drummond and Rambaut 2007). Por otra parte, los métodos bayesianos permiten estimaciones de incertidumbre de los modelos empleados (Drummond *et al.* 2012). Utilizando un enfoque bayesiano MCMC coalescente, los resultados de estos estudios revelaron una tasa media de la evolución de HVR1 en la población de cuasiespecies intra-hospedero del HCV de $4,80 \times 10^{-2}$ sustituciones de aminoácidos/sitio/año (tabla de la figura 37).

Lo que concuerda con estudios recientes, que establecieron una tasa significativamente alta de la evolución de la región HVR1 dentro de un único hospedero. Por otra parte, estos estudios demostraron que la evolución de esta región dentro de un único hospedero resultó ser considerablemente mayor que la de entre diferentes hospederos (Gray *et al.* 2011).

Estudios recientes sobre la evolución de HCV durante la infección crónica, como el caso del paciente donde se aisló esta población de cuasiespecies (Kamila Caraballo-Cortés *et al.* 2013), revelaron que durante la progresión a la cronicidad el

virus evoluciona bajo presión de selección negativa (Campo *et al.* 2008). Estos estudios sugieren que durante el curso de la enfermedad, el virus se desarrolla a través de cuatro etapas: (a) el HCV establece una población en un nuevo hospedero en virtud de selección en ausencia de respuesta inmune específica, (b) la variabilidad de los virus aumenta en la población y variantes de escape inmunitario se generan; (c) la población viral se diversifica en un conjunto de sub-poblaciones que reemplazan la población dominante inicial, y (d) El HCV se asienta bajo una fuerte selección negativa (Ramachandran *et al.* 2011). Esto está de acuerdo con los resultados de este trabajo, ya que se observó una diversificación fuerte y rápida en tres sub-poblaciones diferentes (Figuras 37 y 38).

Curiosamente, mientras que en un trabajo anterior sobre la población de cuasiespecies se habla de una variante principal, que está rodeada por un espectro de mutantes con una cierta distribución de probabilidad (Domingo *et al.* 2012) y la variante principal se percibe a menudo como la variante con la más alta aptitud bajo condiciones ambientales particulares (Jackowiak *et al.* 2014), los resultados de este trabajo, basado en el análisis de la secuencia de ultra-profunda de una población de cuasiespecies, revelaron que este no es el caso en la población de cuasiespecies aquí analizado.

7.2.2 Análisis del árbol filogenético de las secuencias de cuasiespecies de la HVR1 de HCV.

La secuencia más abundante (Seq1, 10,49% de abundancia) está evolucionando en una sola sub-población (Figuras 37 y 38), que muestra una relación genética más estrecha con los haplotipos de la subpoblación 1, y una relación genética más distante con las sub-poblaciones 2 y 3 (Figuras 37 y 38). Estos resultados apoyan el concepto de cuasiespecie que asume que es la población en su conjunto, y no una variante individual, la que es el objetivo real de la selección (Domingo *et al.* 2012). Por lo tanto, la ventaja de un conjunto heterogéneo de variantes reside en la capacidad mejorada de la población de HCV para explorar de manera eficiente un amplia espacio de secuencia para encontrar el que presente el óptimo fitness (Jackowiak *et al.* 2014).

7.2.3 Análisis de redes de mediana de la población de cuasiespecies de la HVR1 de HCV.

El análisis de redes de Unión Media mostró que la población de cuasiespecies estaba compuesta por tres distintas sub-poblaciones (1-3) (Fig. 39). Una vez más, al menos dos sub-poblaciones diferentes (2 y 3), que representa más del 50 % de los haplotipos observados se presentaron en la población de cuasiespecies y están situados en diferentes espacios de la red en comparación con la sub-población 1, que contiene la secuencia principal (Figura 39).

Estos resultados sugieren que una población de cuasiespecies de virus divergentes ocupa grandes extensiones del espacio de secuencias (Escobar-Gutierrez *et al.* 2013).

7.2.4 Co-evolución de sitios de aminoácidos en la HVR1 de las cuasiespecies de HCV.

La detección de co-evolución entre aminoácidos en una proteína mediante el análisis comparativo de secuencias de genes homólogos es una importante fuente para la caracterización funcional y/o estructural de las proteínas (Poon *et al.* 2008). La extensa variabilidad del HVR1 sigue siendo un obstáculo para el desarrollo de estrategias apropiadas de antivirales y el desarrollo de vacunas contra HCV. La HVR1 juega un papel importante en la mediación de entrada a la célula del virus, la neutralización mediada por anticuerpos, y la evasión inmune (Drummond *et al.* 2006). Se ha sugerido que existe una interacción compleja entre HVR1 y los receptores SR-BI y CD81 (Bankwitz *et al.* 2010). Curiosamente, hemos identificado dos posiciones de aminoácidos que juegan un papel clave en la co-evolución dentro de la HVR1 de esta población de cuasiespecies (posiciones 12 y 13), (Figura 40).

Estos aminoácidos interactúan con las posiciones de aminoácidos clave identificados recientemente a ser fundamental para la unión de la proteína de la envoltura al receptor SR-BI del HCV (posiciones 14, 15 y 26) (Drummond *et al.* 2006).

Aunque nuestro modelo *in silico* representa una aproximación para resolver en detalle la estructura 3D de la HVR1 y de la proteína E2 completa (la que por desgracia no está disponible en este momento), nuestros resultados sugieren que la mayoría de estos aminoácidos se encuentran en una estructura de α -hélice ordenada y unas

combinaciones adecuadas de co-evolución entre los aminoácidos puede ser necesaria para mantener esta estructura (Figura 41). Esto está de acuerdo con resultados recientes que revelan que la delección de estos residuos claves situados en el microdominio 1 de HVR1 podría crear cambios conformacionales locales, lo que sugiere que su papel en la interacción con el SR-BI puede ser indirecta (Drummond *et al.* 2006).

8. Objetivo general B

8.1 Análisis de miRNAs asociados al HCV.

Los miRNAs codificados por virus (vsRNAs) son únicos, ya que regulan no sólo su propia expresión génica, sino también la expresión génica del hospedero (Sullivan y Ganem 2005).

A los vsRNAs se les puede clasificar en 2 grupos:

a) los que son análogos de miRNAs celulares, un subconjunto de miRNAs virales han evolucionado para imitar a los miRNA celulares, se les conoce como "análogos". Son vsRNAs que comparten la región semillas con miRNAs celulares, se ha demostrado que regulan las transcripciones a través de los mismos sitios blancos que sus miRNAs celulares homólogos. Es de suponer que este tipo de redes reguladoras evolucionan para efectuar funciones específicas, por ejemplo, la inhibición de la apoptosis.

Se estima que <10% de los miRNAs codificadas por virus humanos actualmente anotados podría imitar miRNAs celulares por poseer secuencias de semillas idénticas.

b) los que son específicos viral.

Aunque más de 250 vsRNAs son conocidos (Sullivan y Ganem 2005), no se posee una clara comprensión de su funcionalidad.

La mayoría de las funciones atribuidas a vsRNAs se pueden agrupar en 3 categorías:

- i) Prolongar la longevidad de las células infectadas.
- ii) Evadir la respuesta inmune.
- iii) Regular genes virales o del hospedero para limitar el ciclo lítico.

Hasta la fecha, miRNAs codificados por virus de RNA se han identificado en el virus de la hepatitis C (HCV) (Shrivastava *et al.* 2015), el virus de inmunodeficiencia humana (VIH) (Zhang *et al.* 2014), el virus de la leucemia bovina (BLV) (Rosewick *et*

al. 2013), el virus del Nilo Occidental (VWN) (Hussain y Asgari 2014), el virus Dengue (DENV) (Ospina-Bedoya *et al.* 2014), y el coronavirus Síndrome Respiratorio de Medio Oriente (MERS) (Hasan *et al.* 2014)..

Muy recientes estudios revelaron que el antigenoma de otro virus RNA citoplasmático, el virus de la hepatitis A (VHA), podrían ser transformados en miRNA por la maquinaria de procesamiento celular (Shi *et al.* 2014).

Con el fin de obtener una idea del posible rol del antigenoma del HCV en relación con la codificación de miRNAs, es que computacionalmente identificamos posibles objetivos de microRNA humanos en el antigenoma de virus de la hepatitis C (HCV) en base a la cepa de referencia H77 (genotipo 1a, NC_004102).

Nuestro estudio puede ayudar a entender mejor la interacción patógeno hospedero, así como contribuir al desarrollo de nuevas terapias antivirales contra el HCV.

8.2. Materiales y Métodos

Para este trabajo, se utilizó la secuencia genómica completa del virus de la hepatitis C (HCV) cepa de referencia H77 (genotipo 1a, NC_004102), que incluía los ORF y las regiones 5' y 3' NCR del RNA viral.

La figura 42 muestra un diagrama de flujo del proceso de predicción computacional empleado

8.2.1. Análisis de secuencias

La secuencia antigenómica completa se obtuvo *in silico* por complemento inverso utilizando el software del programa MEGA (Tamura *et al.* 2011).

El programa VMir (Grundhoff 2011), se utilizó para analizar las secuencias de la cepa H77 del HCV. VMir es un programa de predicción específicamente diseñado para identificar pre-miRNA en los genomas virales.

Con este enfoque, se analizaron las posibles estructuras de horquilla pre-miRNA, utilizando, utilizando filtrado de los parámetros (tamaño horquilla mínima de 60 nucleótidos (nt), tamaño máximo de la horquilla de 120 nt, puntuación mínima horquilla de 115. Sólo se consideraron para los análisis posteriores las estructuras en horquilla de los candidatos a pre-miRNA con un puntaje mínimo de 140.

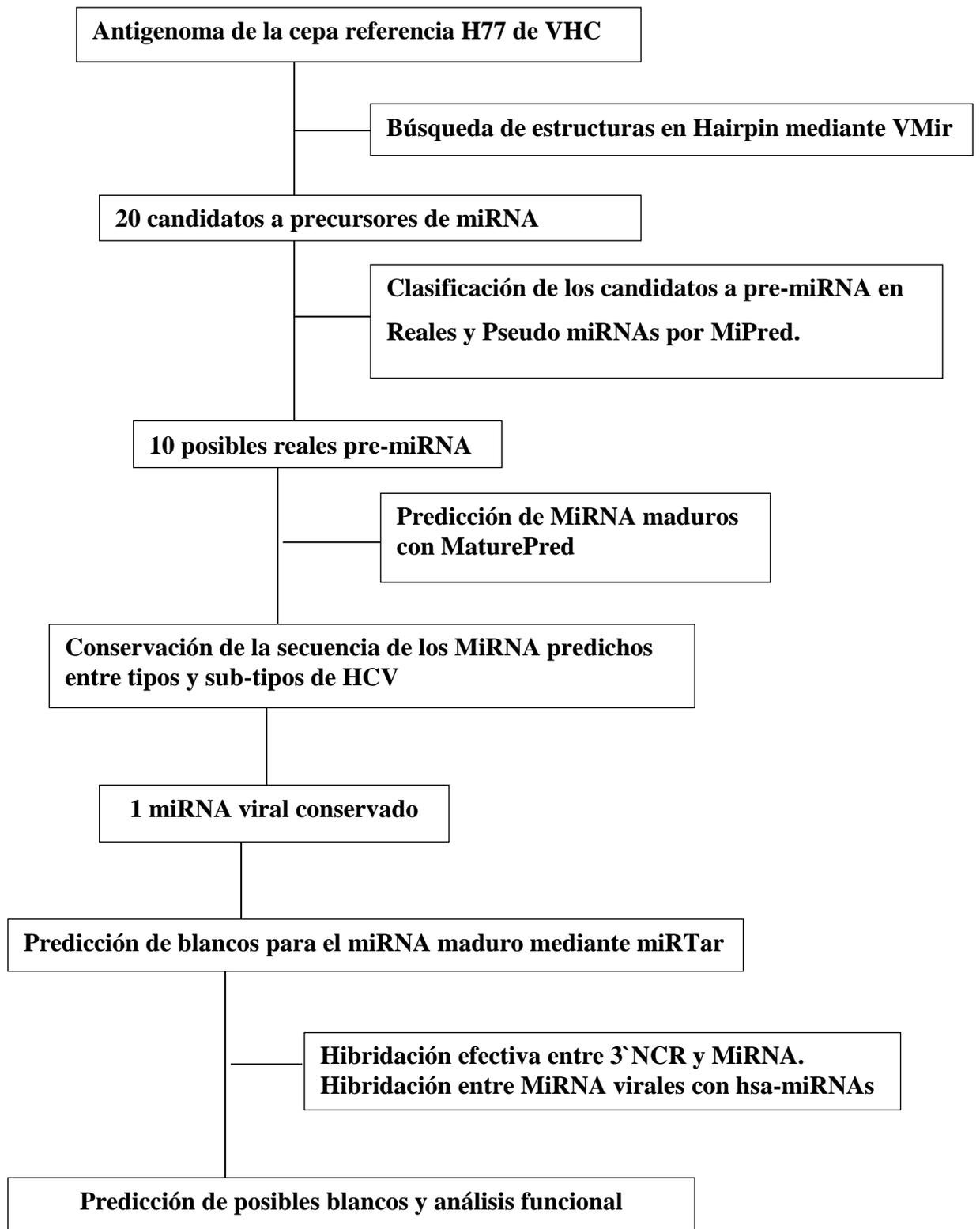


Figura 42. Esquema representando la metodología empleada para la predicción computacional de los miRNAs virales.

8.2.2. Confirmación de las secuencias de los posibles pre-miRNA

Para discriminar verdaderos pre-miRNAs de otras estructuras en horquilla (pseudo horquillas) se empleó MiPred (Jiang *et al.* 2007).

MiPred (disponible en: <http://www.bioinf.seu.edu.cn/miRNA/>), decide si cada horquilla es un verdadero pre-miRNA o bien un pseudo pre-miRNA.

8.2.3. Hibridación de los Pre-miRNA con miRNAs humanos.

Se buscaron la similitud nucleotídica de cada una de las secuencias de los candidatos pre-miRNA clasificados por MiPred como verdaderos pre-miRNA con todos los microRNAs humanos mediante el menú BÚSQUEDA de la base de datos miRBase (Kozomara y Griffiths-Jones 2014). (disponible en: (<http://www.mirbase.org/search.shtml>)). La hibridación entre miRNA humanos y los precursores miRNAs se determinó mediante la herramienta RNAhybrid (Kruger y Rehmsmeier 2006), (disponible en: <http://bibiserv2.cebitec.uni-bielefeld.de/RNAhybrid/>).

RNAhybrid es una herramienta para encontrar la hibridación con mínima energía libre (MFE), la que se fijó en -10 kcal/mol entre una cadena larga y una cadena corta de RNA, es ampliamente utilizado para la predicción de miRNA.

8.2.4. Identificación de secuencias de miRNA maduros.

Con el propósito de extraer miRNA maduro (miRNA*) a partir de horquillas de pre-miRNA, hemos empleado la herramienta Web MaturePred (disponible en: <http://nclab.hit.edu.cn/maturepred/>).

Este software utiliza un modelo basado en Support Vector Machine (SVM) que predice la posición inicial de un miRNA mediante la realización de análisis discriminante en contra de la estructura en horquilla a analizar, utilizando como conjunto de entrenamiento diversas características conocidas de los real/pseudo dúplex miRNA/miRNA*, tales como: posicionamiento característico, funciones relacionadas con la energía, características relacionadas con la estructura y función, y características relacionadas con la estabilidad.

8.2.5. Conservación de los miARN maduros entre genotipos y subtipos de HCV

Con el fin de ver cuan conservadas están las secuencias de los miRNA predichos en la cepa H77 de HCV (genotipo 1a) entre todos los tipos y subtipos de HCV aislados se utilizo una alineación curada que incluye las cepas referencia de todos los tipos y subtipos de las HCV, obtenido de la Base de Datos del HCV (disponible en: <http://hcv.lanl.gov/content/sequence/NEWALIGN/align.html>).

8.2.6. Predicción de la estructura secundaria de los precursores miRNA

El servidor web RNAfold (Lorenz *et al.* 2011), (disponible en <http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) se utilizó para predecir la estructura secundaria de los pre-miRNAs. En todos los casos, las estructuras plegables con centroide fueron representados.

8.2.7. Hibridación entre virales pre-miRNAs y miRNAs humanos

Hibridación efectiva entre miRNA humanos y los precursores miRNAs se determinó mediante la herramienta RNAhybrid (Kruger y Rehmsmeier 2006), (disponible en: <http://bibiserv2.cebitec.uni-bielefeld.de/RNAhybrid/>). RNAhybrid es una herramienta para encontrar la hibridación con mínima energía libre (MFE), de una cadena larga y una cadena corta de RNA, es ampliamente utilizado para la predicción de miRNA.

Para la selección de potenciales miRNA con mínimo de energía libre se fijó como valor umbral -10 kcal/mol.

8.2.8. Predicción de posibles objetivos y análisis funcional de enriquecimiento.

Con el fin de identificar las relaciones entre miRNAs predichos y sus genes blancos, empleamos miRTar (disponibles en: <http://mirtar.mbc.nctu.edu.tw>), (Hsu *et al.* 2011). Los parámetros para la hibridación se establecieron en un valor de -14 kcal/mol de mínima energía libre (MFE) y una alineación con score ≥ 140 .

Determinamos una relación entre los miRNAs predichos en la región 3' NCR, con transcriptos génicos que participan en seis vías metabólicas diferentes (la vía de

apoptosis, la vía de señalización de quimioquinas, la vía de interacción citoquina/receptor de citoquinas, vía de señalización Jak-STAT, vía señalización mTOR y vía de señalización del receptor de células T).

Con el fin de enriquecer los potenciales blancos se analizaron mediante mapas de vía KEGG de miRTar para elucidar las funciones biológicas de miRNAs en dichas vías biológicas (Hsu *et al.* 2011). Sólo los genes con un valor de $p \leq 0,05$ se incluyeron en el análisis.

8.2.9. Hybridization entre los genes blancos transcriptos y los miRNA maduros de la 3'NCR del HCV.

Para confirmar una hibridación eficaz entre los genes blancos y los miRNA de HCV predichos, hemos empleado la herramienta RNAhybrid (disponible en: <http://bibiserv2.cebitec.uni-bielefeld.de/rnahybrid/>), (Kruger y Rehmsmeier 2006).

RNAhybrid es una herramienta para la búsqueda de la energía libre mínima de hibridación entre una cadena larga y una cadena corta de RNA, ampliamente utilizada para la predicción de genes blancos de miRNA. El MFE se fijó en -14 kcal/mol para todas las hibridaciones.

8.3. Resultados

La predicción computacional representa una estrategia ampliamente utilizada y efectiva para identificar nuevos miRNAs que posteriormente se pueden examinar y validar por enfoques experimentales (Thirugnanasambantham *et al.* 2013). El antigenoma de los virus de RNA está presente en el citoplasma y forma diferentes estructuras secundarias que constituyen potenciales pre-miRNAs.

8.3.1. Predicción de posibles pre-miRNAs virales en el antigenoma

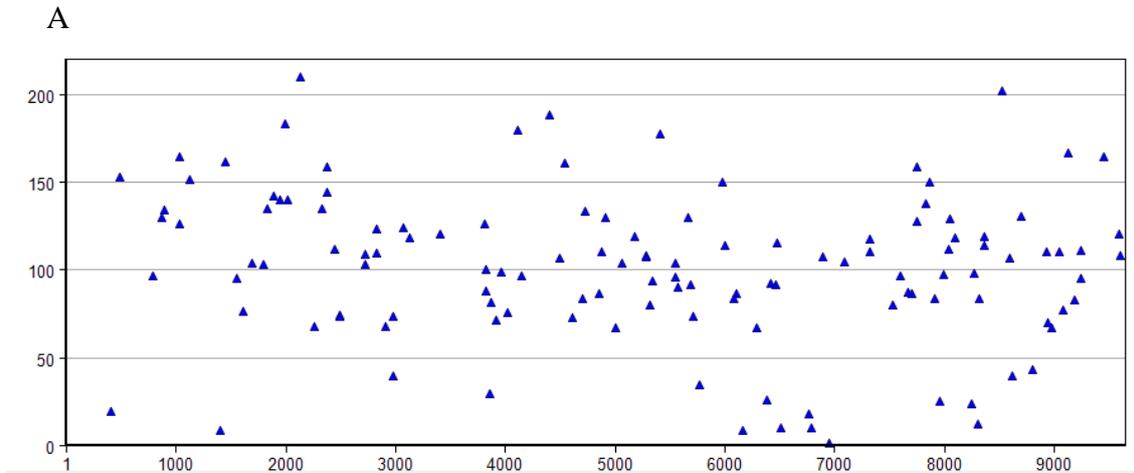
Con el fin de observar si el antigenoma HCV podría ser doblado en estructuras tallo-bucle similares a pre-miRNA, se analiza su supuesta capacidad de codificación de miRNAs utilizando el programa Vmir (Grundhoff 2011), la secuencia completa del antigenoma de la cepa de referencia HCV H77 fue seleccionado. Cuando se utilizan los parametros predeterminados, se identificaron 251 estructuras en horquillas candidatos

(Figura 43a). Para obtener las estructuras en horquilla más fiables a pre-miRNA, filtramos la salida de VMir utilizando la configuración personalizada, estableciendo un valor de corte de 60 nucleótidos (nt) para el tamaño mínimo de la horquilla, 120 nt de tamaño máximo, y 115 puntaje mínimo de la horquilla. Estos ajustes revelaron 82 horquillas candidatas a pre-miRNA (Figura 43b).

De estos 82 potenciales pre-miRNAs, se seleccionaron las posibles estructuras de tallo-bucle con puntajes más altos de 140, para su posterior análisis (20 estructuras). Con el fin de confirmar la presencia de estas posibles estructuras pre-miRNA y clasificarlos en pre-miRNAs reales o pseudo, se utilizó MiPred (Jiang *et al.* 2007).

Diez diferentes estructuras pre-miRNA identificados por VMir, denominadas MD11, MD36, MD19, MD18, MD3, MD39, MD30, MD1, MD26 y MD14, se clasificaron por MiPred como reales pre-miRNAs. Estas estructuras están presentes en diferentes regiones genómicas (Tabla 2). Para cada una de las secuencias identificadas como real pre-miRNA se buscaron su similitud nucleotídica con todos los microRNAs humanos mediante el uso de filtro para miRNA humanos en el menú de búsqueda de la base de datos miRBase (Kozomara y Griffiths-Jones 2014).

Como se muestra en la Tabla 3, diez secuencias identificadas como candidato pre-miRNA tienen una similitud de secuencia significativa con los miRNAs humanos, hsa-miR-671-3p, hsa-miR-668-3p, hsa-miR-3118, hsa-miR-3621, hsa-miR-615-5p, hsa-miR-1293, hsa-miR-664b, hsa-miR-4436b-5p, hsa-miR-3150-5p y hsa-miR-6880.



Posición respecto al antigenoma de la cepa H77 de HCV

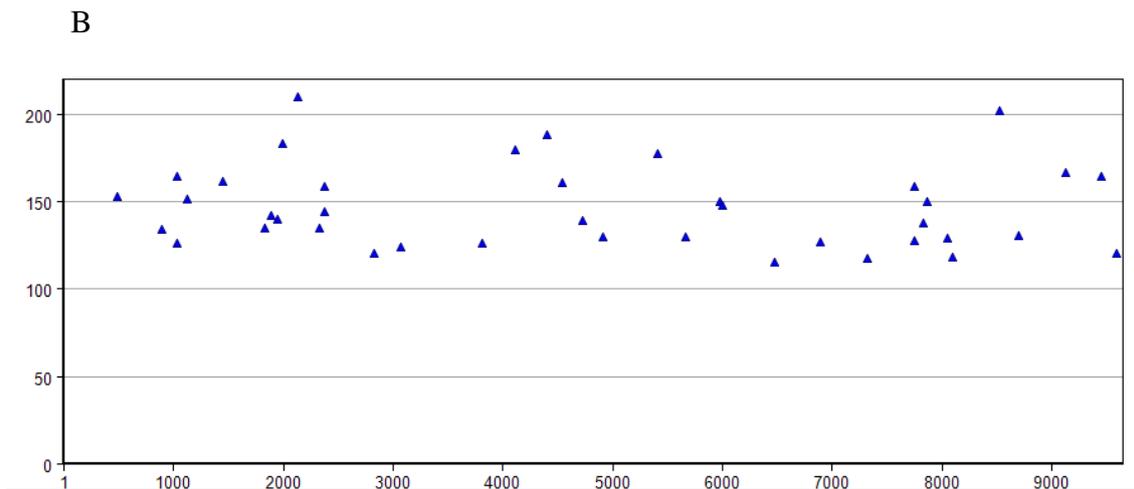


Figura 43. Gráficos de las salidas del análisis del antigenoma de la cepa H77 de HCV obtenidos con VMir Viewer.

En el eje “X” se representa la posición respecto al antigenoma de la cepa H77 y en el eje “Y” el score de los respectivos hairpins.

En (A) se muestran todos los candidatos a pre-miRNAs obtenidos con los valores por defecto de los parámetros de búsqueda.

En (B) se muestran los candidatos a pre-miRNAs obtenidos utilizando como parámetros de búsqueda: tamaño mínimo de hairpin 60 nt, tamaño máximo de hairpin 120 nt, y score mínimo de hairpin 115.

8.3.2. Predicción de miRNAs maduros.

Con el fin de predecir las secuencias de los miRNA maduros en las estructuras pre-miRNAs identificados en el antigenoma de cepa VHC H77, utilizamos MaturePred [25]. Al igual que otros virus de RNA estudiados anteriormente, los miRNAs de la cepa H77 del HCV puede estar ubicado en cualquiera de los dos brazos en la estructura secundaria de la horquilla. De los diez miRNAs identificados en el antigenoma de HCV cepa H77, cuatro están situados en el extremo 5' del brazo de la estructura en horquilla, mientras que seis están en el extremo 3' del brazo (Tabla 4).

8.3.3. Conservación de las secuencias de miRNAs maduros entre los difentes tipos y sub-tipos de HCV.

EL HCV presenta un alto grado de variabilidad genética (Smith *et al.* 2014). La alta tasa de error de la RNA polimerasa dependiente de RNA y la presión ejercida por el sistema inmune del hospedero, ha impulsado la evolución de HCV en 7 genotipos diferentes y un número creciente de subtipos. Con el fin de observar si las secuencias de los miRNA maduro presentes en los miRNAs predichos en el antigenoma de la cepa H77 del HCV se conservan entre diferentes tipos y subtipos de HCV, se realizó una búsqueda de la similitud entre las secuencias incluidas en la base de datos de HCV que incluye todos los tipos y subtipos (Kuiken *et al.* 2005), mediante la utilización del programa MEGA (Tamura *et al.* 2011).

Los resultados de estos estudios revelaron que todas las secuencias de miRNA maduros presentes en las 10 estructuras pre-miRNA reales, encontrados en el antigenoma de HCV de la cepa H77 se conservan sólo entre cepas del sub-genotipo 1a del HCV.

Curiosamente, secuencia de miRNA madura presente en pre-miRNA-MD39 presenta 100% de similitud entre las cepas de HCV que pertenecen a los genotipos 1a, 1b, 1c, 2a, 2c, 2k, 5, 6 y 7, y 95% de similitud con el genotipo 2a, 3a, 3b, 3k y 4.

Esto puso de manifiesto un alto grado de conservación de la secuencia de este miRNA maduro entre las cepas de HCV. Con el fin de observar si también la estructura pre-miRNA se conserva entre todos los genotipos del HCV, se predijo la estructura

secundaria de pre-miRNA-MD39 usando secuencias correspondientes de pre-miRNA-MD39 de diferentes tipos y subtipos de HCV, (Figura 44b).

Como se puede ver en la figura, estructuras secundarias muy similares se obtienen utilizando cepas de HCV de todos los tipos. Esto reveló que esta estructura se conserva entre los genotipos del HCV.

8.3.4. Predicción de potenciales blancos para el miRNA predicho en el antígenoma del HCV y su anotación funcional.

Comprender la dinámica entre miRNAs virales y sus blancos es muy importante para entender la complejidad de la regulación biológica y la interacción virus-hospedero. La predicción *in silico* de los genes blancos de miARN proporciona un enfoque adecuado para la identificación de posibles sitios blancos en función de su complementariedad completa o parcial con los miRNAs.

La búsqueda de blancos de la secuencia miARN-MD39 madura implicada en seis diferentes vías metabólicas se realizó por medio de la utilización de la base de datos miRTar (Hsu *et al.* 2011). Observamos 17 blancos para el predicho HCV-miRNA-MD39 (tabla 5). Estos blancos están implicados principalmente en la apoptosis así como en vías de las respuestas inmunes de la célula hospedera (tabla 5).

La Gene-ontología es una herramienta útil para la extracción de datos de genes y sus anotaciones funcionales. Con el fin de profundizar en estos temas, se realizó un análisis de enriquecimiento funcional utilizando mapas de vías de KEGG de la base de datos miRTar (Hsu *et al.* 2011). Un total de 38 genes se identificaron (Figura 45).

8.3.5. Hibridación eficaz de miARN predicho en la cepa H77 de HCV y 3' NCR objetivos de transcripción de genes.

Para reconfirmar hibridaciones eficaces entre los blancos de los MiRNAs identificados en 3' NCR por miRTar (Hsu *et al.* 2011) y el HCV-miR-MD39, observamos sus patrones de hibridación y se calculó el mínimo de energía libre de cada hibridación por medio de la utilización de la herramienta de RNAhybrid. Hemos utilizado un MFE de -14 kcal/mol.

La tabla 6, muestra las hibridaciones eficaces que se encontraron en todos los casos incluidos en este análisis.

Tabla 2. Alineamiento entre los candidatos a pre-miRNAs con hsa-miRNAs.

Posición ^a del Hairpin	Región ^b	Escore ^c	Alineamiento entre hsa-miRNA y pre-miRNA				
MD11	2086-2179	NS5A	209.6	UserSeq	25	uccggaucccccaggcuccccc	45
				hsa-miR-671-3p	1	uccgguucucaggcuccacc	21
MD36	8468-8563	E1	201.7	UserSeq	30	guagagggccgagcagaggg	11
				hsa-miR-668-3p	4	guagugggcccagccgagug	23
MD19	4340-4458	NS3	188.2	UserSeq	60	gugacuucauucugaaca	77
				hsa-miR-3118	2	gugacugcauuugaaaa	19
MD18	4062-4146	NS4B	179.5	UserSeq	22	gcgggacgaggucugcagg	40
				hsa-miR-3621	2	gcgggucgggucugcagg	20
MD3	968-1086	NS5B	164.8	UserSeq	19	gggggucccccggg	32
				hsa-miR-615-5p	1	gggggucccccggug	14
MD39	9406-9490	5'-UTR	164.2	UserSeq	25	gccccaaucuccaggc	10
				hsa-miR-1293	7	gcacaaucuccagac	22
MD30	7705-7783	E2	158.5	UserSeq	4	ucauuugcaccaccagc	19
				hsa-miR-664b-3p	2	ucauuugcaccaccagc	17
MD1	432-533	NS5B	152.9	UserSeq	50	ccacauauggcagccugcc	69
				hsa-miR-4436b-5p	3	ccacuucugccugccugcc	22
MD26	5952-6053	NS3	148.1	UserSeq	6	aaccuugaggagcggccagc	26
				hsa-miR-3150b-5p	2	aaccucgaggaucccccagc	22
MD14	2348-2414	NS5A	144.5	UserSeq	2	uggaggugguagegggcagc	21
				hsa-miR-6880-5p	1	ugguggaggaagagggcagc	20

^a Posición relativa con respecto al antígeno de la cepa H77 de HCV. ^b Región del genoma de HCV. ^c Escore obtenido por VMir

Table 3. Hibridación de hsa-miRNAs con pre-miRNAs predichos del antigenoma de HCV.

Harpin	Human miRNA	Hybridization	MFE ^a (kcal/mol)	Pairing pattern
MD11	hsa-miR-671-3p	target 5' C G U U 3' GUC CUGAGA CCGGA CGG GACUCU GGCCU miRNA 3' CACCU U 5'	-27.0	5' seed
MD36	hsa-miR-668-3p	target 5' G A AGA C 3' UAG GGGCCGAGC GGGUGGCG AUC CCCGGCUCG CUCACUGU miRNA 3' A G 5'	-40.1	5' canonical
MD19	hsa-miR-3118	target 5' G AUUU G U 3' UCAUGAUGU GGU AU AGUAUUACG UCA UG miRNA 3' CUUAAA G U 5'	-17.4	5' seed
MD18	hsa-miR-3621	target 5' U AGG GGGCCU U 3' CUGCAGG CCGA UCUGC GACGUCU GGCU GGGCG miRNA 3' GG C 5'	-25.3	5' canonical
MD3	hsa-miR-615-5p	target 5' U UG G C 3' UCUG GU UGGGGG UCCC AGGC CG GCCCCU GGGG miRNA 3' U U UG G 5'	-33.7	5' canonical

Table 4. Secuencias de los posibles MiRNA maduros a partir del antigenoma de la cepa H77 de HCV

Hairpin	Posición ^a	Región ^b	Score ^c	5' Brazo Maduro	Brazo 3'Maduro
MD11	2086-2179	NS5A	209.6	AUCCGGAUCCCCAGGCUCCCC	AAGAAUAGGACUCAACGUCGGA
MD36	8468-8563	E1	201.7	GCUCCCGACAAGCAGAUCGAUA	GGGGAGUUUGCCGUCCCUGGUG
MD19	4340-4458	E2	188.2	GUGAUUGGGUGCGUCAGGGUGA	UGUAUAGCAGGGGUGUUGGCC
MD18	4062-4146	NS4B	179.5	GGGACGCGGUCUCGAGGAGGCC	UGAACUGCUCAGCGAGCAUCAU
MD3	968-1086	NS5B	164.8	GGCGGAGUACCUUGGUCAUAGCC	GCCGCGUCCUCCUGGACCCCCG
MD39	9406-9490	5'-UTR	164.2	UCCAGGCAUUGAGCGGGUUUUAU	CAAUCCGGUGUACUCACCG
MD30	7705-7783	E2	158.5	CUGUAGGUAGGCGCGCCGACC	CCACGGGGCUGGGAGUGAAGC
MD1	432-533	NS5B	152.9	UUGAAGAGGUACUUGCCACUA	CUCUGGACAGAAGCCUAGCGCG
MD26	5952-6053	NS3	148.1	AAGGUCUUGGUCCACAUUGGUA	CUUGGGUGAUGCGAUGGUCCUC
MD14	2348-2414	NS5A	144.5	UAGCGGGCAGCCAUGGACCACA	AGGCUUUUCCACGUCUCUAC

^a Posición relativa con respecto al antigenoma de la cepa H77 de HCV. ^b Región del genoma de HCV. ^c Score de VMir. La secuencia en negrita corresponde a la secuencia del MiRNA maduro.

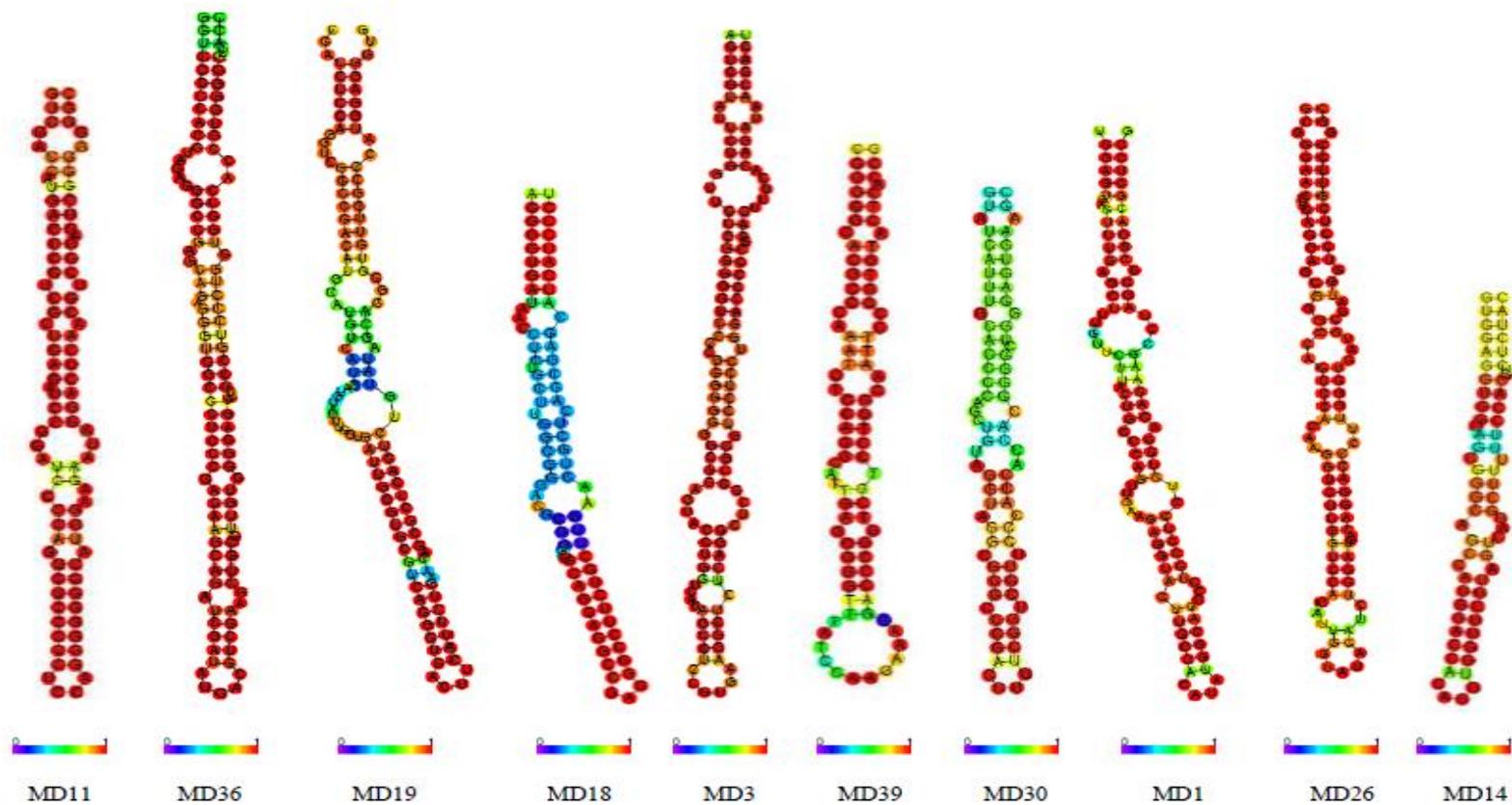


Figura 44a.- Predicción de la estructura secundaria de las horquillas potenciales candidatos a pre-miRNAs

Se muestran las estructuras secundarias predichas de los 10 pre-miRNAs, identificados por su nombre en la parte inferior de la figura. Las barras en la parte inferior de las estructuras denotan las probabilidades de los apareamientos entre bases pares de bases. Sólo fueron representadas las estructuras con centroide. El mínimo de energía libre (MFE) encontrado para las estructuras de los pre-miARN fueron los siguientes: MD11 = -48,40 kcal/mol; MD 36 = -66,80 kcal/mol; MD19 = -54,00 kcal/mol; MD18 = -43,30 kcal/mol; MD3 = -66,00 kcal/mol; MD39 = -35,70 kcal/mol; MD30 = -31,60 kcal/mol; MD1 = -39,00; MD26 = -44,40 kcal/mol y MD14 = -23,90 kcal/mol

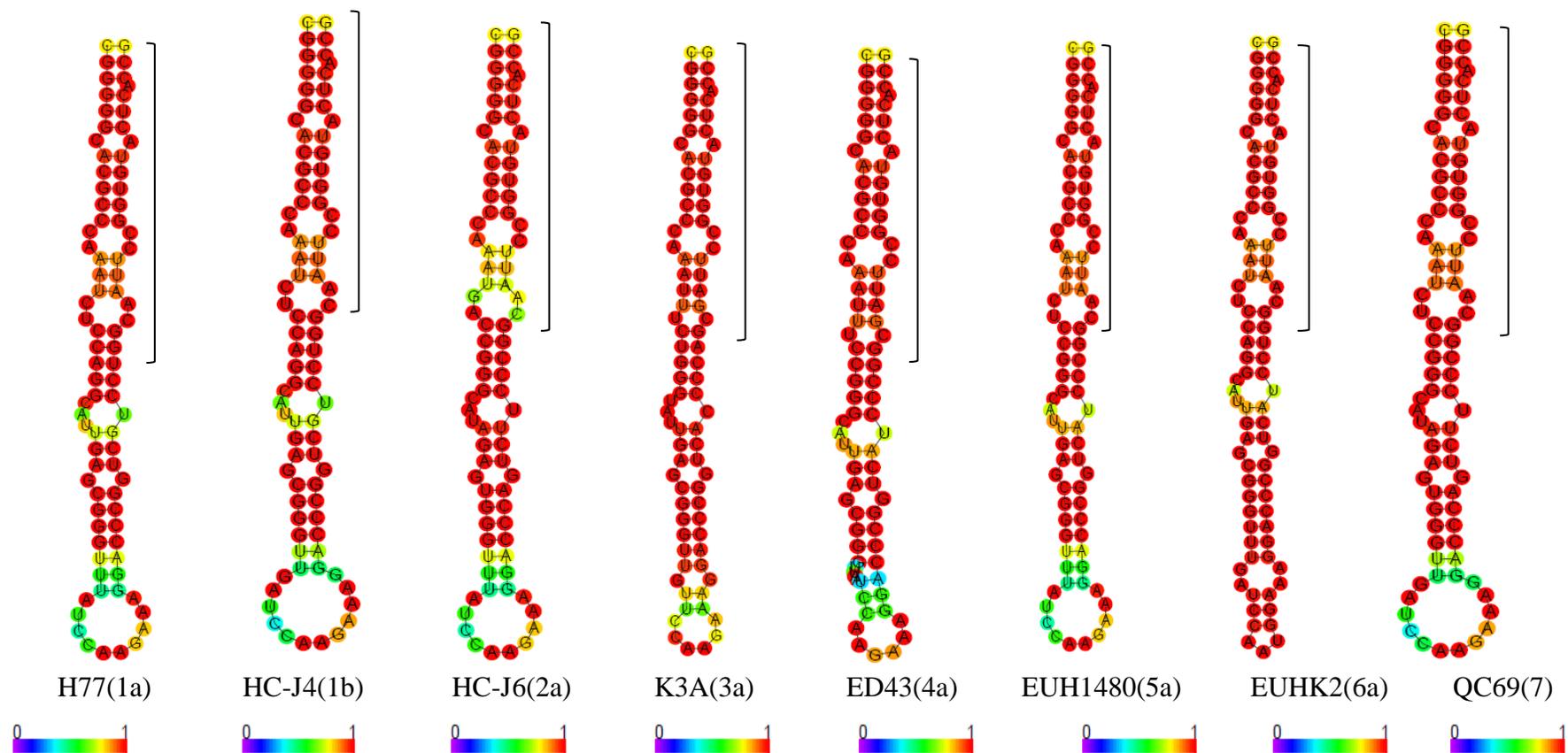


Figura 44b.- Predicción de la estructura secundaria del pre-miRNA-MD39 en los diferentes genotipos del HCV.

Se muestran las estructuras secundarias predichas para el pre-miARN-MD39 utilizando las secuencias de diferentes genotipos y subtipos del HCV. Las barras en la parte inferior de las estructuras indican probabilidades de apareamiento entre las bases. Sólo se representan las estructuras con centroide. Las cepas del HCV se indican por su nombre, y sus genotipos se muestran entre paréntesis junto al nombre de la cepa. Las secuencias correspondientes al miARN maduro se indican con un corchete junto a cada estructura. La energía libre mínima (MFE) encontrada para las estructuras de los pre-miARN fueron las siguientes: H77 (1a) = -35,70 kcal/mol; HC-J4 (1b) = -35,60 kcal/mol; HC-J6 (2a) = -37,60 kcal/mol; K3A (3a) = -39,90 kcal/mol; ED43 (4a) = -39,90 kcal/mol; EUH1480 (5a) = -37,90 kcal/mol.

Table 5. Predicción de blancos del HCV-miR-MD39-3p identificados por análisis *in silico*.

Hairpin	Predicted miRNA	Metabolic pathway	Targeted proteins	Protein description
MD39	HCV-miR-MD39-3p	Apoptosis	XIAP	X-linked inhibitor of apoptosis
		Chemokine signaling pathway	GNAI3 ITK	guanine nucleotide binding protein IL2-inducible T-cell kinase
		Cytokine-cytokine receptor	LIFR IL21R	leukemia inhibitory factor receptor alpha interleukin 21 receptor
		Jak-STAT signaling pathway	CREBBP LIFR OSM PIAS3 IL21R	CREB binding protein leukemia inhibitory factor receptor alpha oncostatin M protein inhibitor of activated STAT, 3 interleukin 21 receptor
		mTOR signaling pathway	PRKAA2 KIAA1303 CAB39L	protein kinase, AMP-activated, alpha2 catalytic subunit Raptor calcium binding protein 39-like
		T cell receptor signaling pathway	CD8A ITK	CD8a molecule IL2-inducible T-cell kinase

ICOS

inducible T-cell co-stimulator

Tabla 6. Hibridación del posible miRNA en 3' NCR de H77 HCV y los blancos identificados mediante análisis *in silico*.

miRNA predicho	Blanco	Nº acceso	Hybridization	MFE ^a (kcal/mol)	Patrón Hibridación
HCV-miRMD39-3p	XIAP	U45880	target 5' U G GGG G 3' CGG GAG GGGAUUG GCC CUC CCUUAAC miRNA 3' A AUGUGG 5'	-19.2	5'canonica
	GNAI3	M27543	target 5' A AAGA CA U C 3' UGGUGAGUA GCC GG UUG GCCACUCAU UGG CC AAC miRNA 3' G UU 5'	-23.1	3'compensatoria
	ITK	D13720	target 5' A AA A 3' GA ACAUUGGAAUUG CU UGUGGCCUUAAC miRNA 3' GCCA CA 5'	-21.2	5'canonica
	LIFR	X61615	target 5' U A UG UA GG A 3' GUGA GUAU U GGA UG CACU CAUG G CCU AC miRNA 3' GC UG UA 5'	-19.2	3'compensatoria
	OSM	AF129855	target 5' U G G A 3' GUGA GUG GCUGGGA CACU CAU UGGCCUU miRNA 3' GC G AAC 5'	-22.6	5'canonica

Tabla 6. Hibridación del posible miRNA en 3' NCR de H77 HCV y los blancos identificados mediante análisis *in silico*. (Cont.).

Mirna predicho	Blanco	Nº acceso	Hybridization	MFE ^a (kcal/mol)	Patrón Hibridación
IL21R	AF254667		target 5' G A A 3' GGUGAG ACAUC CCACUC UGUGG miRNA 3' G A CCUUAAC 5'	-21.9	3'compensatoria
CREBBP	NM_004380		target 5' G A GCUGGG UG UUG G 3' CGG GAG GC CGG UG GCC CUC UG GCC AC miRNA 3' A A UG UUA 5'	-22.3	3'compensatoria
PIAS3	AB021868		target 5' U CAG CCU C 3' GG GAGU GC CCGG AUUG CC CUCA UG GGCC UAAC miRNA 3' G A U U 5'	-21.2	3'compensatoria
PRKAA2	BC069823		target 5' G UUGGCA U 3' GGUG GU ACACC CCAC CA UGUGG miRNA 3' G U CCUUAAC 5'	-17.9	3'compensatoria

Table 6. Hibridación del posible miRNA en 3' NCR de H77 HCV y los blancos identificados mediante análisis *in silico*. (Cont.).

Mirna predicho	Blanco	Nº acceso	Hybridization	MFE ^a (kcal/mol)	Patrón Hibridación
KIAA1303	NM_020639	target 5' U GG GGG ACACUGGAA CC CUC UGUGGCCUU miRNA 3' G A A AAC 5'	A 3'	-24.4	5'canonica
CAB39L	AK022639	target 5' G A C 3' GGUGAGUG GGGG CCACUCAU CCUU miRNA 3' G GUGG AAC 5'		-22.7	3'compensatoria
CD8A	NM_001768	target 5' U CUUA C A 3' GGUGAG AC CUGGAA CCACUC UG GGCCUU miRNA 3' G A U AAC 5'		-23.4	3'compensatoria
ICOS	AB023135	target 5' U A AAACAU A 3' GGU GG UGC CCGGAAUUG CCA UC AUG GGCCUUAAC miRNA 3' G C U 5'		-23.7	5'canonica

^a mínima energía libre



Figura 45.- Análisis de enriquecimiento funcional de los blancos previstos para el HCV-miR-MD39.

Las vías metabólicas se muestran en la parte izquierda de la figura. Las barras son proporcionales al número de genes/términos en el análisis.

* Significa un valor de $p < 0,05$, ** significa un valor de $p < 0,01$

8.4. Discusión

Los miRNAs desempeñan papeles críticos en muchos procesos biológicos, tales como el crecimiento celular, la diferenciación de tejidos, la proliferación celular, el desarrollo embrionario, y la apoptosis (Hussain y Asgari 2014; Wang y Chang 2011). Estudios recientes revelaron que genomas de virus de diferentes familias codifican miRNAs (Kincaid y Sullivan 2012).

Los miRNAs virales han sido identificados tanto por estrategia de clonación tradicional a partir de células infectadas por virus (Pfeffer *et al.* 2004), como por predicción computacional (Cui *et al.* 2006). Esta última estrategia ha sido utilizada para identificar miRNAs en numerosos virus (Besecker *et al.* 2009).

Utilizando el programa VMir, se ha confirmado que varios genomas de virus son capaces de codificar miRNAs, lo que sugiere que el programa VMir es una herramienta eficaz para la búsqueda de nuevos miRNAs virales (Seo *et al.* 2009; Shi *et al.* 2014). En este estudio, mediante el uso de enfoques bioinformáticos, 10 estructuras pre-miRNAs con altos puntajes en VMir y clasificados como verdaderos pre-miRNA utilizando MiPred fueron identificados en el antigenoma del HCV cepa de referencia H77 (Tabla 2).

Esto está de acuerdo con resultados recientes que identificaron nuevos miRNAs en el antigenoma de otro virus hepático, el virus de Hepatitis A (Shi *et al.* 2014). Por otra parte, en un estudio reciente llevado a cabo en virus Dengue se encontraron todos los miRNAs predichos en la hebra inversa del RNAm viral (Ospina-Bedoya *et al.* 2014).

Las secuencias de los miRNA maduros de estos 10 miRNAs potenciales se conservan entre el del genotipo 1a del HCV, al que pertenece la cepa H77. Sin embargo, sólo las secuencias maduras del miRNA del pre-miRNA-MD39 resultaron ser conservadas entre todos los genotipos y subtipos del HCV.

Por otra parte, las predicciones de las estructuras secundarias de pre-miRNA-MD39 utilizando secuencias correspondientes de todos los genotipos y subtipos del VHC revelaron estructuras adecuadas y similares (Figura 44b)

La 5' NCR del genoma viral tiene una doble función en el ciclo de replicación del HCV; primero en la cadena positiva al funcionar como un sitio interno de entrada al ribosoma (IRES), capaz de conducir la traducción del RNA, y la síntesis de la

poliproteína; y segundo, en la cadena negativa, proporcionando elementos de replicación actuando en cis (CRES). Se ha demostrado que la 5' NCR y el extremo 3' complementario del RNA de cadena negativa es capaz de adoptar diferentes estructuras secundarias, de acuerdo con sus diferentes funciones en la traducción y la replicación del RNA (Smith *et al.* 2002). El extremo 3' de la cadena negativa contiene dos dominios diferentes: el dominio I, que abarca los nucleótidos -230 a -1 y prevé que sea una región estructurada, y el dominio II, una región menos estructurada que contiene las secuencias complementarias al IRES del 5' NCR de cadena positiva (Schuster 2002).

El mapeo del pre-miARN predicho en el extremo 3' de cadena negativa, abarca los nucleótidos -241 a -157.

Con el fin de profundizar en la posibilidad de observar si las secuencias de cadena negativa 3' NCR pueden sostener una estructura como la predicha para el pre-miRNA, modelamos las secuencias 3'terminal de la cadena negativa de la cepa H77 de HCV (nucleótidos -250 a -1) por medio de la utilización del programa MFold .

Mientras que seis diferentes estructuras se pueden obtener para esta región genómica usando MFold, con MFEs que van desde -116 hasta -112 kcal/mol, uno de ellos, con un MFE de -112 kcal/mol, predice la presencia de la horquilla pre-miRNA en el extremo 3' de la cadena negativa de la cepa H77 de HCV (Figura 46). Este resultado sugiere que el extremo 3' de la cadena negativa del RNA de HCV podría tener la capacidad de adoptar esta estructura bajo ciertas condiciones.

Los miRNAs juegan un papel muy importante en el establecimiento de la infección por HCV, apuntando a factores celulares del hospedero, necesarios para aumentar la replicación y el crecimiento del HCV. La alteración de la expresión de los miRNAs está involucrado en la patogénesis del hígado asociada a la infección por el HCV (Gong *et al.* 2015).

Estudios recientes revelaron que la expresión alterada de miRNAs está implicada en la patogénesis asociada con la infección por HCV mediante el control de las vías de señalización tales como respuesta inmune, la proliferación y la apoptosis (Shrivastava *et al.* 2013). Esto está de acuerdo con los resultados encontrados en este trabajo, ya que los blancos previstos para el HCV-miR-MD39 son miembros de las vías metabólicas relacionadas con la apoptosis y la respuesta inmune (Tabla 5 y Figura 4).

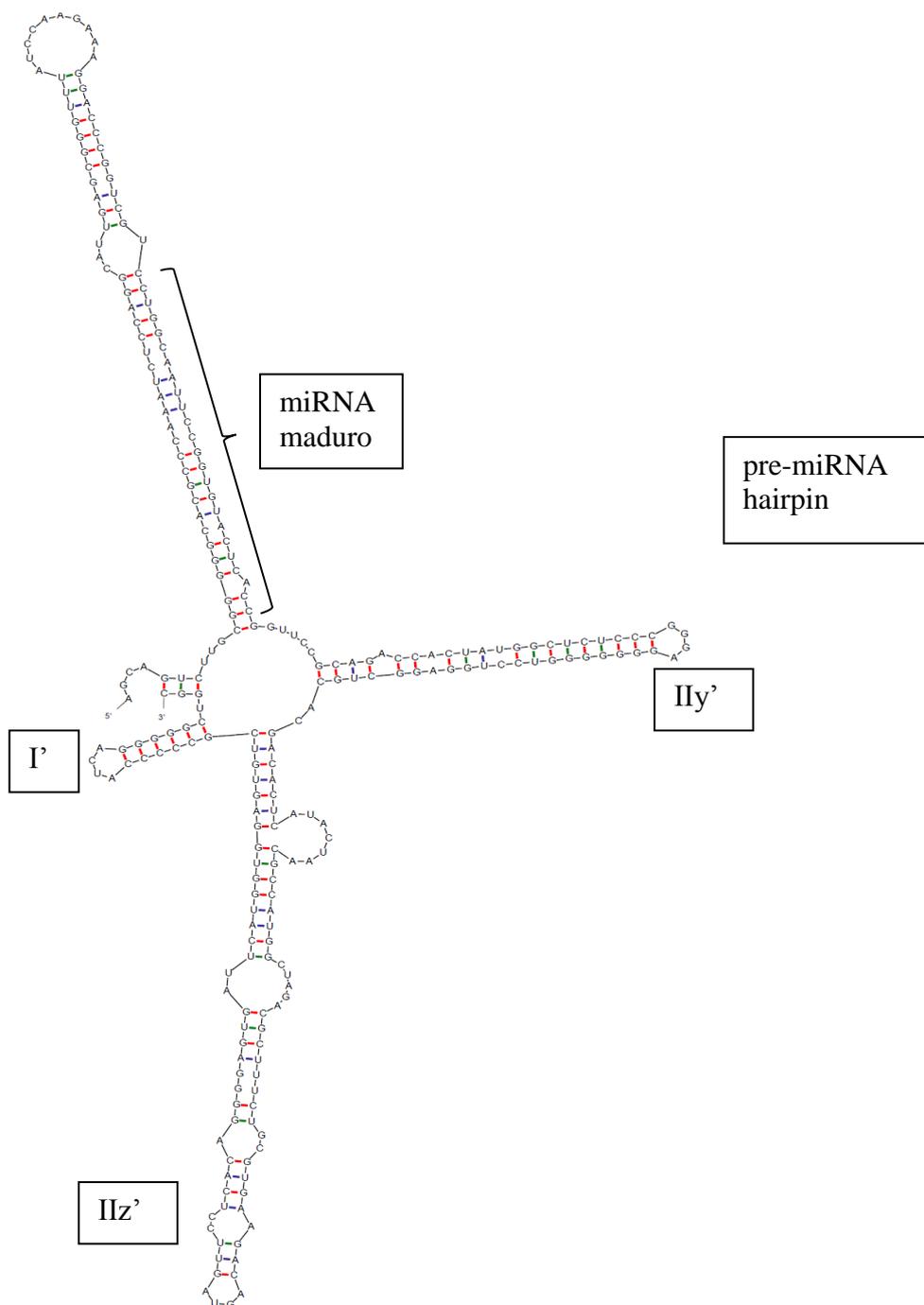


Figura 46.- Predicción de la estructura secundaria del extremo 3' de la cadena negativa del RNA de la cepa H77 de HCV.

Una estructura secundaria de las secuencias 3'-terminal de la cadena negativa de la cepa H77 de HCV, que abarca las posiciones de nucleótidos -250 a -1, se muestra. Un MFE de -112 kcal/mol se obtuvo para la estructura mostrada. El pre-miRNA predicho en estos estudios se muestra en la parte superior de la figura. Los dominios de la estructura secundaria descritos previamente por Smith *et al.* (Smith, *et al.*,2002) se muestran por su nombre al lado de los dominios.

Por otra parte, se observaron hibridaciones eficaces entre los 3' NCR y secuencias maduras de HCV-miR-MD39 en todos los casos (Tabla 6).

Se cree que la lesión hepática en la infección por HCV es causada por las respuestas inmunes del hospedero, y no por efectos citopáticos virales.

La infección por HCV inhibe la translocación nuclear de NF-kappa B y la expresión de proteínas anti-apoptóticas NF-kappa B-dependientes, tales como el inhibidor ligado al cromosoma X de la proteína apoptótica (XIAP), y la disminución de los niveles del ARN mensajero de XIAP y de la proteína XIAP se ha observado en hígados con hepatitis C crónica (Park *et al.* 2012). Esto está de acuerdo con los resultados encontrados en este trabajo, ya que XIAP fue identificado como un blanco del miARN predicho (Tabla 5).

Aunque XIAP es una molécula anti-apoptótica, también está involucrada en muchas otras vías, incluyendo la regulación de la inmunidad innata (Aguilar and Latour 2015). Por otra parte, los miembros de la familia de las proteínas inhibidoras de apoptosis (como XIAP) su función no está restringida a la inhibición de la apoptosis, estas proteínas tienen un espectro mucho más amplio de acción, lo que sugiere su implicación en la regulación de las respuestas inflamatorias. La inflamación está altamente regulada por ubiquitinación informes recientes revelan el papel de estas proteínas, incluyendo XIAP, como ligasas de ubiquitina que regulan la inmunidad innata y la inflamación (Estornes *et al.* 2015).

EL HCV también ha desarrollado estrategias para bloquear la vía de respuesta del interferón (IFN). Durante la infección por HCV, proteínas virales están implicadas como reguladores negativos de la vía de respuesta del IFN a través de la obstrucción de la vía de señalización JAK/STAT (Yang *et al.* 2015).

Curiosamente, los miembros de la vía de señalización JAK/STAT también han sido identificados como blancos para el VHC-miR-MD39 (Tabla 5). La vía mTOR es un controlador de la traducción celular y la síntesis de macromoléculas (Mannova *et al.* 2005).

Estudios recientes demostraron que el HCV podría activar esta vía mediante TSC1/TSC2, un regulador negativo de mTOR. De esta forma el HCV promueve la expresión mTOR, lo que resulta en la activación de la traducción y la síntesis de proteínas. Curiosamente, la proteína-quinasa PRKAA2, un miembro de la misma vía de señalización TSC1/TSC2, se encontró que era un blanco para HCV-miARN-MD39 (Tabla 5).

En general, este estudio apoya la idea de que el antigenoma de un virus RNA citoplasmático puede codificar de forma natural miRNA funcionales, lo que está de acuerdo con los últimos resultados obtenidos para otros virus citoplasmáticos. Los resultados de este trabajo revelaron un candidato a miRNA viral, que debe ser confirmado por análisis experimental.

9. Conclusiones

1) La implementación de análisis de redes, como complemento de los análisis filogenéticos bayesianos, permite una forma clara de visualizar las variaciones evolutivas dentro de la cuasiespecie, Las redes filogenéticas podrían pues desempeñar un importante papel en la reconstrucción de la historia evolutiva, dado que permiten visualizar mejor las relaciones entre los haplotipos circulantes en la nube de las cuasiespecies y facilitan el análisis de los modelos evolutivos implicados.

Los resultados de este trabajo ponen de manifiesto la idoneidad de estos análisis hacia la importante meta de poder controlar la infección por HCV.

Es por ello que consideramos a los análisis de redes como un elemento complementario de gran importancia para el estudio de las cuasiespecies.

2) Los resultados de este trabajo ponen de relieve la importancia de los genomas minoritarios en la historia de la población y la evolución del HCV, así como el rol de las nubes mutantes como reservorios de variantes fenotípicas y genéticas para mejorar la capacidad de adaptación del virus. Este estudio revela la importancia de los genomas memorias en la transmisión de la infección, y en la evasión a las defensas inmunitarias del hospedero.

El futuro del tratamiento del HCV se basará en nuestra capacidad de evaluar genomas memoria con mutaciones pre-existentes de resistencia al inicio del estudio y durante el tratamiento, con el fin de decidir el tratamiento combinado antiviral más adecuado. Los resultados de este trabajo ponen de manifiesto la idoneidad de estos análisis hacia la importante meta de poder controlar la infección por HCV.

3) Este estudio nos ha permitido observar la composición, complejidad, y amplitud del espectro de mutante de una población de cuasiespecies de HCV por secuenciación ultra-profunda. La ventaja de analizar un amplio espectro de mutantes nos muestra la capacidad de la población viral para encontrar las porciones del espacio de secuencia para mejorar su fitness.

4) Un enfoque NGS, combinado con un método de reconstrucción de datos adecuada junto con un enfoque de análisis Bayesiano coalescente ha permitido una reconstrucción fiable de la historia y evolución de la población de cuasiespecies del HCV circulante en los pacientes. Lo que tendrán importantes implicaciones para evaluar la presencia de genomas memoria en relación con el tratamiento y el control de la infección por HCV.

5) Mediante la utilización de una serie de herramientas bioinformáticas se revela la posibilidad de la presencia de miRNAs virales en la cadena de RNA de sentido negativo, utilizado como plantilla para la replicación del genoma del HCV.

Mediante la utilización de una serie de herramientas bioinformáticas hemos identificado un miARN presente en el antigenoma de cepa H77 de HCV.

Este miARN mapea en la región 5'NCR del genoma del HCV y se encontró que esta conservado entre los genotipos y subtipos del HCV.

La predicción *in silico* genera 17 genes celulares blancos. Estos potenciales blancos están involucrados en la apoptosis, así como en las vías de respuesta inmune, lo que sugiere que podrían desempeñar un papel en la patogénesis causada por la infección viral.

Además, los resultados de estos estudios revelaron la presencia de un miRNA viral en la cadena de RNA de sentido negativo, utilizada como una plantilla para la replicación del genoma de HCV, tal como se observó para otros virus de RNA

Esta predicción *in silico* es una guía útil para el diseño experimental con el fin de lograr la validación biológica.

Tres de los potenciales miRNAs humanos revelados en este trabajo, hsa-miR-671-3p, hsa-miR-615-5p y hsa-miR-1293, están involucrados en diferentes procesos biológicos importantes, lo que sugiere que podrían desempeñar un papel en la interacción virus-hospedero, durante la infección.

Para ayudar a entender mejor que miRNAs virales son las más relevantes, sería útil contar con un estudio más profundo de los virus que codifican miRNAs.

Además, nuestros resultados proporcionan una base para posteriores trabajos para evaluar las funciones de los genes miRNA codificado a partir del antigenoma de HCV durante la infección por el virus y las interacciones virus-hospedero.

10. Perspectivas

Determinar si en cultivo celulares se constata la expresión de los miRNAs virales predichos por el modelo bioinformático.

Mediante la construcción de plásmidos de expresión de los respectivos candidatos de miRNAs virales, transfectando líneas celulares adecuadas con células Dicer deficientes y con Dicer no deficientes como control negativo.

Detectando la aparición de los miRNAs maduros mediante RT-PCR.

11. Agradecimientos

A mi mujer por el apoyo incondicional durante el desarrollo de esta Tesis.

A mis hijos para que le sirva de ejemplo de no entregarse jamás.

A mi hermana por darme siempre para adelante.

Al Dr. Juan Cristina profe, compañero y amigo por todo lo que me ayudo para poder concretarla.

A.A. mis hermanos.

A la Agencia Nacional de Investigación e Innovación que hizo posible la realización de esta tesis a través del apoyo económico mediante una beca de posgrado nacional.

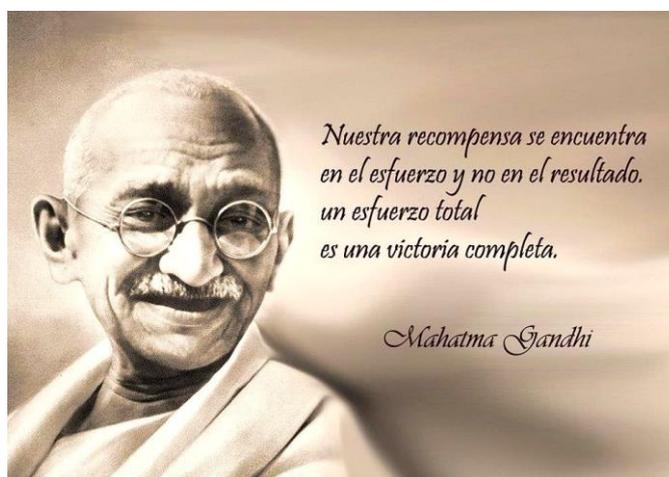
Al Laboratorio de Virología Molecular.

A mis compañeros de ayer y de hoy.

A Alvaro Fajardo por sus consejos y apoyo durante todo este tiempo.

A los docentes integrantes del tribunal por sus aportes.

A todos los que de una forma u otra han hecho posible la realización de esta tesis.



12. Bibliografía

- **Abe H., Hayes C.N., Hiraga N., Imamura M., Tsuge M. et al. (2013).** A translational study of the resistance emergence using sequential direct-acting antiviral agents for hepatitis C using ultra-deep sequencing. *American Journal of Gastroenterology* **108**: 1464-1472.
- **Abergel A., Ughetto S., Dubost S., Bonny C., Aublet-Cuvelier B., Delarocque-Astagneau E., Bailly J.L., Bommelaer G., Casanova S., Delteil J., Deny P., Laurichesse H., Odent-Malaure H., Roussel J., Peigue-Lafeuille H., Henquell C. (2007).** The epidemiology and virology of hepatitis C virus genotype 5 in central France. *Aliment. Pharmacol. Ther.* **26**: 1437-1446.
- **Afdhal N.H., McHutchison J.G., Zeuzem S., et al. (2011).** Pharmacogenetics and Hepatitis C Meeting Participants. Hepatitis C pharmacogenetics: state of the art in 2010. *Hepatology* **53(1)**: 336–345.
- **Aguilar C., and Latour S. (2015).** X-linked Inhibitor of Apoptosis Protein Deficiency: More than an X-linked Lymphoproliferative Syndrome. *Journal of Clinical Immunology* **35**: 331-338, 2015.
- **Airaksinen A., Pariente N., Menendez-Arias L., Domingo E. (2003).** Curing of foot-and-mouth disease virus from persistently infected cells by ribavirin involves enhanced mutagenesis. *Virology* **311**: 339–349.
- **Alter H.J., and Seeff L.B. (2000).** Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. *Semin Liver Dis* **20(1)**:17–35.
- **Alter M.J. (2007)** Epidemiology of hepatitis C virus infection. *World Journal of Gastroenterology* **13**: 2436–2441.
- **Amadei B., Urbani S., Cazaly A., et al. (2010).** Activation of natural killer cells during acute infection with hepatitis C virus. *Gastroenterology*. **138**: 1536–1545.
- **Bandelt H.J. and Dress A.W.M. (1992).** A canonical decomposition theory for metrics on a finite set. *Adv. Math.* **92**: 47–105.
- **Bandelt H.J., Forster P., Röhl A. (1999).** Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution* **16**: 37–48.
- **Bankwitz D., Steinmann E., Bitzegeio J., Ciesek S., Friesland M., et al. (2010).** Hepatitis C virus hypervariable region 1 modulates receptor interactions, conceals

the CD81-binding site, and protects conserved neutralizing epitopes. *Journal of Virology* **84**: 5751–5763.

- **Bartenschlager R., Lohmann V., Penin F. (2013).** The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nat. Rev. Microbiol.* **11**: 482-496.
- **Bartosch B., Verney G., Dreux M., Donot P., Morice Y., Penin F., Pawlotsky J.M., Lavillette D., Cosset, F.L. (2005).** An interplay between hypervariable region 1 of the hepatitis C virus E2 glycoprotein, the scavenger receptor BI, and high-density lipoprotein promotes both enhancement of infection and protection against neutralizing antibodies. *J. Virol.* **79**: 8217–8229.
- **Bartosch B., Vitelli A., Granier C., Goujon C., Dubuisson J., Pascale S., Scarselli E., Cortese R., Nicosia A., Cosset F.L. (2003).** Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J. Biol. Chem.* **278**: 41624–41630.
- **Barzon L., Lavezzo E., Militello V., Toppo S., Palu G. (2011).** Applications of next-generation sequencing technologies to diagnostic virology. *International Journal of Molecular Sciences* **12**: 7861–7884.
- **Beerenwinkel N. (2011).** Ultra-deep sequencing for the analysis of viral populations. *Current Opinion in Virology* **1**: 413–418.
- **Besecker M.I., Harden M.E., Li G., Wang X.J., Griffiths A. (2009).** Discovery of herpes B virus-encoded microRNAs. *J. Virol.* **83**:3413–3416.
- **Bittar C., Jardim A.C., Yamasaki L.H., de Queiroz A.T., Carareto C.M., et al. (2010).** Genetic diversity of NS5A protein from hepatitis C virus genotype 3a and its relationship to therapy response. *BMC Infect. Dis.* **10**: 36.
- **Bittar C., Jardim A.C.G., Yamasaki L.H.T., Carareto C.M.A., Pinho J.R.R., et al. (2013).** On Hepatitis C Virus Evolution: The Interaction between Virus and Host towards Treatment Outcome. *PLoS ONE* **8(4)**: e62393.
- **Bostan N., Mahmood T. (2010).** An overview about hepatitis C: a devastating virus. *Crit. Rev. Microbiol.* **36(2)**: 91-133.
- **Bouckaert R. (2010).** DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* **26**: 1372-1373.

- **Brady M.T., McDonald A.J., Rowan A.G., Mills K.H., (2003).** Hepatitis C virus non structural protein 4 suppresses Th1 responses by stimulating IL-10 production from monocytes. *Eur. J. Immunol.* **33**: 3448-3457.
- **Branch A.D., Stump D.D., Gutierrez J.A., Eng F., Walewski J.L. (2005).** The hepatitis C virus alternate reading frame (ARF) and its family of novel products: the alternate reading frame protein/Fprotein, the double-frameshift protein, and others. *Semin. Liver Dis.* **25**: 105-117.
- **Brimacombe C.L., Grove J., Meredith L.W., Hu K., Syder A.J., Flores M.V., Timpe J.M., Krieger S.E., Baumert T.F., et al. (2011).** Neutralizing antibody-resistant hepatitis C virus cell to cell transmission. *J. Virol.* **85**: 596–605.
- **Briones C., Domingo E. (2008).** Minority report: hidden memory genomes in HIV-1 quasispecies and possible clinical implications. *AIDS Rev.* **10**: 93–109.
- **Bryant D. and Moulton V. (2004).** NeighborNet: an agglomerative algorithm for the construction of planar phylogenetic networks. *Mol. Biol. Evol.* **21**: 255–265.
- **Cabot B., Martell M., Esteban J.I., Sauleda S., Otero T., Esteban R., Guardia J., Gomez J. (2000).** Nucleotide and amino acid complexity of hepatitis C virus quasispecies in serum and liver. *J. Virol.* **74**: 805–811.
- **Caillet-Saguy C., Lim S.P., Shi P.Y., Lescar J., Bressanelli S. (2014).** Polymerases of hepatitis C viruses and flaviviruses: structural and mechanistic insights and drug development. *Antiviral Res.* **105**: 8-16.
- **Campo D.S., Dimitrova Z., Mitchell R.J., Lara J., Khudyakov Y. (2008).** Coordinated evolution of the hepatitis C virus. *Proceedings of the National Academy of Science USA*, **105**: 9685–9690.
- **Charleston M.A. (1998).** Jungles: a new solution to the host/parasite phylogeny reconciliation problem. *Math. Biosci.* **149**: 191–223.
- **Chatterji U., Lim P., Bobardt M.D., Wieland S., Cordek D.G., Vuagniaux G., Chisari F., Cameron C.E., Targett-Adams P., Parkinson T., et al. (2010).** HCV resistance to cyclosporin A does not correlate with a resistance of the NS5A-cyclophilin A interaction to cyclophilin inhibitors. *J. Hepatol.* **53**(1):50-56.
- **Chen M., Sallberg M., Sonnerborg A., et al. (1999).** Limited humoral immunity in hepatitis C virus infection. *Gastroenterology.* **116**(1): 135–143.

- **Choo Q.L., Kuo G., Weiner A.J., Overby L.R., Bradley D.W., Houghton M. (1989).** Isolation of a Cdna clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**: 359-62.
- **Clayton R.F., Owsianka A., Aitken J., Graham S., Bhella D., Patel A.H. (2002).** Analysis of antigenicity and topology of E2 glycoprotein present on recombinant hepatitis C virus-like particles. *J.Virol.* **76**: 7672–7682.
- **Comas I., Moya A., Gonzalez-Candelas F. (2005).** Validating viral quasispecies with digital organisms: a reexamination of the critical mutation rate. *BMC Evol. Biol.* **5**: 5.
- **Contreras A.M., Hisasa Y., He W., Terella A., Schmidt E.V., Chung R.T. (2002).** Viral RNA mutations are region specific and increased by ribavirin in a full-length hepatitis C virus replication system. *J. Virol.* **76**: 8505–8517.
- **Cook G.A, Opella S.J. (2011).** Secondary structure, dynamics, and architecture of the p7 membrane protein from hepatitis C virus by NMR spectroscopy. *Biochim. Biophys. Acta.* **1808(6)**:1448-53.
- **Cox A.L., Mosbrugger T., Lauer G.M., et al. (2005).** Comprehensive analyses of CD8+ T cell responses Turing longitudinal study of acute human hepatitis C. *Hepatology.* **42(1)**: 104–112.
- **Cristina J., Lopez F., Moratorio G., López L., Vasquez S., García-Aguirre L., Chunga A. (2005).** Hepatitis C virus F protein sequence reveals a lack of functional constraints and a variable pattern of amino acid substitution. *J. Gen. Virol.* **.86**: 115-120.
- **Crotty S., Cameron C.E., Andino R., 2001.** RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl. Acad. Sci. U.S.A.* **98**: 6895–6900.
- **Cui C., Griffiths A., Li G., Silva L.M., Kramer M.F., Gaasterland T., Wang X.J., Coen D.M. (2006).** Prediction and identification of herpes simplex virus 1-encoded microRNAs. *J. Virol.* **80**:5499–5508.
- **Cuthbert J.A. (1994).** Hepatitis C: progress and problems. *Clin. Microbiol. Rev.* **7(4)**:505-32.
- **Degasperi E., Aghemo A. (2014).** Sofosbuvir for the treatment of chronic hepatitis C: between current evidence and future perspectives. *Hepat. Med.* **29**: 25-33.
- **Del Portillo A., et al. (2011).** Multiploid inheritance of HIV-1 during cell-to-cell infection. *J. Virol.* **85**: 7169 –7176.

- **Delang L., Coelmont L., Neyts J. (2010).** Antiviral therapy for hepatitis C virus: beyond the standard of care. *Viruses* **2**(4):826-866.
- **Delport W., Poon A.F., Frost S.D., Kosakovsky-Pond S.L. (2010).** Datamonkey 2010 a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, **26**: 2455-2457.
- **Demetriou V.L., Kyriakou E., Kostrikis L.G. (2011).** Near-full genome characterisation of two natural intergenotypic 2k/1b recombinant Hepatitis C virus isolates. *Adv. Virol.* **2011**: 710438.
- **Di Bisceglie A.M. (1998).** Hepatitis about hepatitis C: a devastating virus. *C. Lancet* 351-355.
- **Di Lorenzo C., Angus A.G., Patel A.H. (2011).** Hepatitis C virus evasion mechanisms from neutralizing antibodies. *Viruses* **3**: 2280–2300.
- **Domingo E. (2000).** Viruses at the edge of adaptation. *Virology.* **270**: 251–253.
- **Domingo E. (2006).** Quasispecies: concepts and implications for virology *Curr. Top. Microbiol. Immunol.* **299**: 1– 422.
- **Domingo E., Martin V., Perales C., Grande-Pérez A., García-Arriaza J., Arias A. (2006).** Viruses as quasispecies: biological implications. *Current Topics in Microbiology and Immunology* **299**: 51-82.
- **Domingo E., Sheldon J., Perales C. (2012).** Viral quasispecies evolution. *Microbiology and Molecular Biology Review* **76**: 159-216.
- **Dowd K.A., Netski D.M., Wang X.H., Cox A.L., Ray S.C. (2009).** Selection pressure from neutralizing antibodies drives sequence evolution during acute infection with hepatitis C virus. *Gastroenterology* **136**: 2377–2386.
- **Drummer H.E., Boo I., Maerz A.L., Pountourios P. (2006).** A conserved Gly436-Trp-Leu-Ala-Gly-Leu-Phe-Tyr motif in hepatitis C virus glycoprotein E2 is a determinant of CD81 binding and viral entry. *J. Virol.* **80**: 7844–7853.
- **Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. (2006).** Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**: e88.
- **Drummond A.J., Rambaut A. (2007).** BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.
- **Drummond A.J., Suchard M.A., Xie D., Rambaut A. (2012).** Bayesian Phylogenetics with BEAUti and the BEAST 1.7 *Molecular Biology and Evolution* **29**: 1969–1973.

- **Drummond, A.J., Rambaut, A., Shapiro, B. and Pybus, O.G. (2005).** Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution* **22**: 1185-1192.
- **Duarte E.A. et al. (1994).** Subclonal components of consensus fitness in an RNA virus clone. *J. Virol.* **68**: 4295– 4301.
- **Durand D.B.V. Halldorsson Vernot B. (2005).** Hybrid micro-macroevolutionary approach to gene tree reconstruction. Pp. 250–264 in S. Miyano, J. Mesirov, S. Kasif, S. Istrail, P. Pevzner, and M. Waterman, eds., *Lecture Notes in Computer Science*. vol. **3500**. Springer-Verlag, Berlin.
- **Duverlie G., Khorsi H., Castelain S., Jaillon O., Izopet J., Lunel F., Eb F., Penin F., Wychowski C. (1998).** Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity. *J. Gen. Virol.* **79**: 1373–1381.
- **Eigen M, Schuster P. (1978).** The hypercycle—a principle of natural self-organization. C. The realistic hypercycle. *Naturwissenschaften* **65**: 341–369.
- **Eigen M. (1971).** Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**: 465–523.
- **El Tayebi H.M., Hosny K.A., Esmat G., Breuhahn K., Abdelaziz A.I. (2012).** miR-615-5p is restrictedly expressed in cirrhotic and cancerous liver tissues and its overexpression alleviates the tumorigenic effects in hepatocellular carcinoma. *FEBS Lett* **586**:3309-3316.
- **El-Shabrawi M.H., Kamal N.M. (2013).** Burden of pediatric hepatitis C. *World J. Gastroenterol.* **19**:7880-7888.
- **Enomoto N., Sakuma I., Asahina Y. et al. (1996).** Mutations in the nonstructural protein 5A and response to interferon in patients with chronic hepatitis C virus 1b infection. *N. Engl. J. Med.* **334**: 77–81.
- **Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T., Yamamoto C., Izumi N., Marumo F., Sato C. (1995).** Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b: sensitivity to interferon is conferred by amino acid substitution in the NS5A region. *J. Clin. Invest.* **96**: 224–230.

- **Erdtmann L., Franck N., Lerat H., Le Seyec J., Gilot D., Cannie I., Gripon P., Hibner U., Guguen-Guillouzo C. (2000).** The Hepatitis C virus NS2 protein is an inhibitor of CIDE-B-induced apoptosis. *J. Biol. Chem.* **278**: 18256-18264.
- **Eriksson N., Pachter L., Mitsuya Y., Rhee S.Y., Wang C., Gharizadeh B., Ronaghi M., Shafer R.W., Beerenwinkel N. (2008).** Viral population estimation using pyrosequencing. *PLoS Comput. Biol.* **4**: e1000074.
- **Escobar-Gutierrez A., Soudeyns H., Larouche A., Carpio-Pedroza J., Martinez-Gaurneros A. et al. (2013).** Vertical transmission of hepatitis C virus: a tale of multiple outcomes. *Infection, Genetics and Evolution*, **20**: 465-470.
- **Estornes Y., Bertrand M.J. (2015).** IAPs, regulators of innate immunity and inflammation. *Seminars in Cellular Development and Biology* **39**: 106-114.
- **European Association for Study of Liver. (2014).** EASL Clinical Practice Guidelines: management of hepatitis C virus infection. *J. Hepatol.* **60(2)**: 392-420.
- **Fang S.H., Hwang L.H., Chen D.S., Chiang B.L. (2000).** Ribavirin enhancement of hepatitis C virus core antigen-specific type 1 T helper cell response correlates with the increased IL-12 level. *J. Hepatol.* **33**: 791-798.
- **Farci P. (2011).** New insights into the HCV quasispecies and compartmentalization. *Semin. Liver Dis.* **31**: 356-374.
- **Farci P., Alter H.J., Shimoda A., Govindarajan S., Cheung L.C., Melpolder J.C., Sacher R.A., Shih J.W., Purcell R.H. (1996).** Hepatitis C virus-associated fulminant hepatic failure. *N. Engl. J. Med.* **335**: 631-634.
- **Farris J.S. (1970).** Methods for computing Wagner trees. *Syst. Zool.* **19**: 83-92.
- **Feinstone S.M., Kapikian A.Z., Purcell R.H., Alter H.J., Holland P.V. (1975).** Transfusion-associated hepatitis not due to viral hepatitis type A or B. *N Engl J Med.* **292 (15)**: 767-70.
- **Fishman S.L., Murray J.M., Eng F.J., Walewski J.L., Morgello S., Branch A.D. (2008).** Molecular and bioinformatic evidence of hepatitis C virus evolution in brain. *J. Infect. Dis.* **197**: 597-607.
- **Flint M., Maidens C., Loomis-Price L.D., Shotton C., Dubuisson J., Monk P., Higginbottom A., Levy S., McKeating J.A. (1999).** Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81. *J. Virol.* **73**: 6235-6244.

- **Forster P., Torroni A., Renfrew C., Röhl A. (2001).** Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Molecular Biology and Evolution*. **18**: 1864-1881.
- **Forton D.M., Karayiannis P., Mahmud N., Taylor-Robinson S.D., Thomas H.C. (2004).** Identification of unique hepatitis C virus quasispecies in the central nervous system and comparative analysis of internal translational efficiency of brain, liver, and serum variants. *J. Virol.* **78**: 5170–5183.
- **Fournillier A., Freida D., Defrance T., et al. (2004).** Analysis of B-lymphocyte differentiation in patients infected with hepatitis C virus. *J. Med. Virol.* **72**(4): 566–574.
- **Frick D.N., Rypma R.S., Lam A.M., Gu B., (2004).** The nonstructural protein 3 protease/helicase requires an intact protease domain to unwind duplex RNA efficiently. *J. Biol. Chem.* **279**: 1269–1280.
- **Fried M.W., Hadziyannis S.J. (2004).** Treatment of chronic hepatitis C infection with peginterferons plus ribavirin. *Semin. Liver Dis.* **24**: 47–54.
- **Gaël C., Ivanyi-Nagy R., Gabus C., Boulant1 S., Lavergne J.P., Penin F., Darlix J.L. (2004).** The hepatitis C virus Core protein is a potent nucleic acid chaperone that directs dimerization of the viral (+) strand RNA in vitro. *Nucleic. Acids. Res.* **32**: 2623-2631.
- **Gale M., Blakely C.M., Kwieciszewski B., Tan, M. Dossett S.L., Tang N.M., Korth M.J., Polyak S.J., Gretch D.R., Katze M.G. (1998).** Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol. Cell. Biol.* **18**: 5208–5218.
- **Gale M., Jorth M.J., Tang N.M., Tan S.L., Hopkins D.A., Dever T.E., Polyak S.J., Gretch D.R., Katze M.G. (1997).** Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. *Virology* **230**: 217–227.
- **Gao G., Buskell Z., Seeff L., Tabor E. (2002).** Drift in the hypervariable region of the hepatitis C virus during 27 years in two patients. *J. Med. Virol.* **68**: 60–67.
- **Gimenez-Barcons M., Franco S., Suarez Y. et al. (2001).** High amino acid variability within the NS5A of hepatitis C virus (HCV) is associated with hepatocellular carcinoma in patients with HCV-1b-related cirrosis. *Hepatology* **34**: 158–167.

- **Gong G., Waris G., Tanveer R., Siddiqui A. (2001).** Human hepatitis C virus NS5A protein alters intracellular calcium levels, induces oxidative stress, and activates STAT-3 and NF-kB. *Proc. Natl. Acad. Sci. U.S.A.* **98**: 9599–9604.
- **Gong J., Xing-Xing H.E., Tian D. (2015).** Emerging role of microRNA in hepatocellular carcinoma. *Oncology Let.* **9**: 1027-1033.
- **Gong P., Xie F., Zhang B., Perkins E.J. (2010).** In silico identification of conserved microRNAs and their target transcripts from expressed sequence tags of three earthworm species. *Comput. Biol. Chem.* **34**:313–319.
- **Gong Y., et al. (2013).** TIMP-1 promotes accumulation of cancer associated fibroblasts and cancer progression. *PLoS One* **8(10)**: e77366.
- **Gouttenoire J., Penin F., Moradpour D. (2010).** Hepatitis C virus nonstructural protein 4B: a journey into unexplored territory. *Rev. Med. Virol.* **20**:117-129.
- **Gray R.R., Parker J., Lemey P., Salemi M., Katzourakis A. et al. (2011).** The mode and tempo of hepatitis C virus evolution within and among hosts. *BMC Evolutionary Biology* **11**: 131.
- **Griffin, S. D. et al., (2003).** The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett.* **535**: 34–38.
- **Grundhoff A. (2011).** Computational prediction of viral miRNAs. In: van Rij RP, editor. Antiviral RNAi: Concepts, Methods, and Applications. *Springer Science+Business Media, LLC*.
- **Guan M., Wang W., Liu X., Tong Y., Liu Y., Ren H., Zhu S., Dubuisson J., Baumert T.F., Zhu Y., Peng H., Aurelian L., Zhao P., Qi Z. (2012).** Three Different Functional Microdomains in the Hepatitis C Virus Hypervariable Region 1 (HVR1) Mediate Entry and Immune Evasion. *Journal of Biological Chemistry* **287**: 35631-35645.
- **Guglietta S., Garbuglia A.R., Pacciani V., Scotta C., Perrone M.P., Laurenti L., et al. (2005).** Positive selection of cytotoxic T lymphocyte escape variants during acute Hepatitis C virus infection. *European Journal of Immunology*, **35**: 2627-2637.
- **Gusfield D., and Bansal V. (2005).** A fundamental decomposition theory for phylogenetic networks and incompatible characters. Pp. 217–232 in S. Miyano, J. Mesirov, S. Kasif, S. Istrail, P. Pevzner, and M. Waterman, eds. *Research in Computational Biology*. Lecture Notes in Computer Science, vol. 3500, Springer-Verlag, Berlin.

- **Hadziyannis S., Papatheodoridis G. (2004).** Ensayos recientes sobre la combinación de Peg-interferón y Ribavirina. *Current Hepatitis Report. Edición en español* **1**: 01-08.
- **Hallett M. and Lagergren J. (2000).** New algorithms for the duplication-loss model. Pp. 138–146 in R. Shamir, S. Miyano, S. Istrail, P. Pevzner, and M. Waterman, eds. Proceedings of the fourth annual international conference on computational molecular biology. ACM Press, New York.
- **Hanson R.M. (2010).** Jmol – a paradigm shift in crystallographic visualization. *Journal of Applied Crystallography* **43**: 1250-1260.
- **Hasan M.M., Akter R., Ullah M.S., Abedin M.J., Ullah G.M., Hossain M.Z. (2014).** A Computational Approach for Predicting Role of Human MicroRNAs in MERS-CoV Genome. *Adv. Bioinformatics*. 967946.
- **Herring B.L., Tsui R., Peddada L., Busch M., Delwart E.L. (2005).** Wide range of quasispecies diversity during primary hepatitis C virus infection. *J. Virol.* **79**:4340–4346.
- **Hoffmann T.W., Duverlie G., Bengrine A. (2012).** MicroRNAs and hepatitis C virus: toward the end of miR-122 supremacy. *J. Virol.* **9**: 109.
- **Hong Z., Cameron C.E. (2002).** Pleiotropic mechanisms of ribavirin antiviral activities. *Prog. Drug. Res.* **59**: 41–69.
- **Hraber P.T., Fischer W., Bruno W.J., Leitner T., Kuiken C. (2006).** Comparative analysis of hepatitis C virus phylogenies from coding and non-coding regions: the 5' untranslated region (UTR) fails to classify subtypes. *Virol. J.* **14**:103.
- **Hsu J.B., Chiu C.M., Hsu S.D., et al. (2011).** miRTar: an integrated system for identifying miRNA-target interactions in human. *BMC Bioinformatics*, 10.1186/1471-2105-12-300.
- **Hsu M., Zhang J., Flint M., Logvinoff C., Cheng-Mayer C., Rice C. M., McKeating J.A. (2003).** Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. *Proc. Natl. Acad. Sci. USA* **100**: 7271–7276.
- **Hu K.Q., Vierling J.M., Redeker A.G. (2001).** Viral, host and interferon-related factors modulating the effect of interferon therapy for hepatitis C virus infection. *J. Viral. Hepat.* **8**: 1–18.

- **Huson D.H., Rupp R., Scornavacca C. (2011).** Phylogenetic networks: concepts, algorithms and applications. *Cambridge University Press, Cambridge, UK.*
- **Hussain M., Asgari S.. (2014).** MicroRNA-like viral small RNA from Dengue virus 2 autoregulates its replication in mosquito cells. *Proc. Natl. Acad. Sci. U S A* **111**: 2746–2751.
- **Hwang S.B., Park K.J., Kim Y.S., Sung Y.C., Lai, M.M., (1997).** Hepatitis C virus NS5B protein is a membrane-associated phosphoprotein with a predominantly perinuclear localization. *Virology.* **227**: 439-446.
- **Inchauspe G., Zebedee S., Lee D.H., Sugitani M., Nasoff M., Prince A.M. (1991).** Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. *Proc. Natl. Acad. Sci. U.S.A.* **88**: 10292–10296.
- **Jackowiak P., Kuls K., Budzko L., Mania A., Figlerowicz M. et al. (2014).** Phylogeny and molecular evolution of hepatitis C virus. *Infection, Genetics and Evolution*, **21**: 67-82.
- **Jacobi M.N., Nordahl M. (2006).** Quasispecies and recombination. *Theor. Popul. Biol.* **70**: 479–485. .
- **Janssen H.L., Reesink H.W., Lawitz E.J., Zeuzem S., Rodriguez-Torres M., Patel K., van der Meer A.J., Patick A.K., Chen A., Zhou Y., Persson R., King B.D., Kauppinen S., Levin A.A., Hodges M.R. (2013).** Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **368**: 18/1685–1694.
- **Jardim A.C., Bittar C., Matos R., Yamasaki L.H., Silva R., Pinho J., Fachini R., et al. (2013).** Analysis of HCV quasispecies dynamic under selective pressure of combined therapy. *BMC Infect. Dis.* **13**: 61.
- **Jiang P., Wu H., Wang W., Ma W., Sun X., Lu Z. (2007).** MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic. Acids. Res.* **35**: W339-W344.
- **Johansson D.X., Voisset C., Tarr A.W., Aung M., Ball J.K., Dubuisson J., Persson M.A. (2007).** Human combinatorial libraries yield rare antibodies that broadly neutralize hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **104**: 16269–16274.
- **Jopling C. (2012).** Liver-specific microRNA-122: biogenesis and function. *RNA. Biol.* **9**: 137–142.

- **Kalinina O., Norder H., Magnius L.O. (2004).** Full-length open reading frame of a recombinant hepatitis C virus strain from St Petersburg: Proposed mechanism for its formation. *J. Gen. Virol.* **85**: 1853–1857.
- **Kalinina O., Norder H., Mukomolov S., Magnius L.O. (2002)** A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J. Virol.* **76**: 4034–4043.
- **Kamal S.M. (2011).** Hepatitis C virus genotype 4 therapy: progress and challenges. *Liver Int.* **1**: 45-52.
- **Kamila Caraballo-Cortés K., Zagordi O., Laskus T., Poski R., Bukowska-Osko I., Pawelczyk A. et al. (2013).** Ultradeep pyrosequencing of Hepatitis C virus hypervariable region 1 in quasispecies analysis. *BioMed Research International* **2013**: 626083.
- **Kato N., Sekiya H., Ootsuyama Y., Nakazawa T., Hijikata M., Ohkoshi S., Shimotohno K. (1993).** Humoral immune response to hypervariable region 1 of the putative envelope glycoprotein (gp70) of hepatitis C virus. *J. Virol.* **67**: 3923–3930.
- **Keck Z.Y., Sung V. M.H., Perkins S., Rowe J., Paul, S., Liang T.J., Lai M.M.C., Fong S.K.H. (2004).** Human monoclonal antibody to hepatitis C virus E1 glycoprotein that blocks virus attachment and viral infectivity. *J. Virol.* **78**: 7257–7263.
- **Kieft J. S., Zhou K., Jubin R., Doudna J. A. (2001).** Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* **7**: 194–206.
- **Kincaid R.P., Sullivan C.S. (2012).** Virus-Encoded microRNAs: An Overview and a Look to the Future. *PloS. Pathog.* **8**: e1003018.
- **Koizumi K., Enomoto N., Kurosaki M., et al. (1995).** Diversity of quasispecies in various disease stages of chronic hepatitis C virus infection and its significance in interferon treatment. *Hepatology* **22(1)**: 30–35.
- **Kozomara A., Griffiths-Jones S (2010).** miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic. Acids. Res.* **39**: D152–D157.
- **Kozomara A., Griffiths-Jones S. (2014).** MiRBase: annotating high confidence microRNAs using deep sequencing data,” *Nucleic. Acids. Res.* **42**: D68–D73.
- **Krekulová L., Reháč V., Riley L.W. (2006).** Structure and functions of hepatitis C virus proteins: 15 years after. *Folia. Microbiol. (Praha).* **51(6)**:665-680.

- **Kren B.T., Wong P.Y., Sarver A., Zhang X., Zeng Y., Steer C.J. (2009).** MicroRNAs identified in highly purified liver-derived mitochondria may play a role in apoptosis. *RNA Biol.* **6**:65–72.
- **Kruger J., Rehmsmeier M. (2006).** RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic. Acids. Res.* **34**:W451–W454.
- **Kruskal J.B. (1956).** On the shortest spanning subtree of the graph and the travelling salesman problem. *Proc. Amer. Math. Soc.* **7**: 48–57.
- **Kuiken C., Yusim K., Boykin L., et al. (2005).** The Los Alamos HCV Sequence Database. *Bioinformatics*, **21**:379-384.
- **Kuo G., Choo Q.L., Alter H.J., Gitnick G.L., Redeker A.G., Purcell R.H., Miyamura T., Dienstag J.L, Alter M.J., Stevens C.E., et al. (1989).** An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* **244**: 362-64.
- **Kurbanov F., Tanaka Y., Avazova D., Khan A., Sugauchi F., Kan N., Kurbanova-Khudayberganova D., Khikmatullaeva A., Musabaev E., Mizokami M. (2008).** Detection of hepatitis C virus natural recombinant RF1_2k/1b strain among intravenous drug users in Uzbekistan. *Hepatol. Res* **38**: 457–464.
- **Lauring A.S., Andino R. (2010).** Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* **6**:e1001005.
- **Lavanchy D. (2011).** Evolving epidemiology of hepatitis C virus. *Clin. Microbiol. Infect.* **17**: 107-115.
- **Lee Y.M., Lin H.J., Chen Y.J., Lee C.M., Wang S.F., Chang K.Y., Chen T.L., Liu H.F., Arthur Chen Y.M. (2010).** Molecular epidemiology of HCV genotypes among injection drug users in Taiwan: Full-length sequences of two new subtype 6w strains and a recombinant form_2b6w. *J. Med. Virol.* **82**: 57–68.
- **Lemon,S.M., Walker,C.M., Alter,M.J., Yi,M. (2007).** Hepatitis C virus. In: Knipe, D.M., Howley, P.M., Eds. *Fields virology*, vol 1. Philadelphia: Lippincott Williams & Wilkings 1253–1304.
- **Liang Y., Kang C.B., Yoon H.S. (2006).** Molecular and Structural Characterization of the Domain 2 of Hepatitis C Virus Non-structural Protein 5A. *Mol. Cell.* **22**: 13-20.
- **Lindenbach B.D. (2013).** Virion assembly and release. *Curr. Top. Microbiol. Immunol.* **369**: 199-218.

- **Lindenbach B.D., Murray C.L., Thiel H.J., Rice C.M. (2013).** Flaviviridae. Knipe D.M., Howley P.M. (eds). *Fields Virology, 6th ed.* 712-742.
- **Linder C.R., and Rieseberg L.H. (2004).** Reconstructing patterns of reticulate evolution in plants. *Am. J. Bot.* **91**: 1700–1708.
- **Lindow M., Kauppinen S. (2012).** Discovering the first microRNA-targeted drug. *J. Cell. Biol.* **199**: 407–412.
- **Liu L., Fisher B.E., Dowd K.A., Astemborski J., Cox A.L., Ray S.C. (2010).** Acceleration of hepatitis C virus envelope evolution in humans is consistent with progressive humoral immune selection during the transition from acute to chronic infection. *J. Virol.* **84**: 5067–5077.
- **Lo S.Y., Selby M.J., Ou J.H. (1996).** Interaction between hepatitis C virus core proteína and E1 envelope protein. *J. Virol.* **70**: 5177-5182.
- **Lok A.S., Gardiner D.F., Lawitz E., Martorell C., Everson G.T., Ghalib R.H., Reindollar R., Rustgi V.K., Wendelburg P., Zhu K., et al. (2010).** Combination therapy with BMS-790052 and BMS-650032 alone or with pegylated interferon and ribavirin (PEGIFN/RBV) results in undetectable HCV RNA through 12 weeks of therapy in HCV genotype 1 null responders. *Hepatology* **52(4)**: 877A.
- **López-Jiménez A.J., Clemente-Casares P., Sabariego R., Llanos-Valero M., Bellón-Echeverría I., Encinar J.A., Kaushik-Basu N., Froeyen M., Mas A. (2014).** Hepatitis C virus polymerase-polymerase contact interface: Significance for virus replication and antiviral design. *Antiviral Res.* **3542**: 115-116.
- **Lorenz R., Bernhart S.H, Höner zu Siederdisen C., Tafer H., Flamm C., Stadler P.F., Hofacker I.L. (2011).** ViennaRNA Package 2.0. *Algorith. Mol. Biol.* **6**:26.
- **Lutchman G., Danehower S., Song B.C., Liang T.J., Hoofnagle J.H., et al. (2007).** Mutation rate of the hepatitis C virus NS5B in patients undergoing treatment with ribavirin monotherapy. *Gastroenterology* **132**: 1757–1766.
- **MacDonald A., Harris M. (2004).** Hepatitis C virus NS5A: tales of a promiscuous protein. *J. Gen. Virol.* **85**: 1502–2485.
- **Macdonald A., Mazaleyrat S., McCormick C., Street A., Burgoyne N.J., Jackson R.M., Cazeaux V., Shelton H., Saksela K., Harris M.I. (2005).** Further studies on hepatitis C virus NS5A–SH3 domain interactions: identification of

residues critical for binding and implications for viral RNA replication and modulation of cell signalling. *J. Gen. Virol.* **86**: 1035–1044.

- **Majumder M., Ghosh A.K., Steele R., Ray R., Ray R.B. (2001).** Hepatitis C virus NS5A physically associates with p53 and regulates p21/waf1 gene expression in a p53-dependent manner. *J. Virol.* **75**: 1401–1407.
- **Mannova P., Beretta L., et al. (2005).** Activation of the N-Ras-PI3K-Akt-mTOR pathway by hepatitis C virus: Control of cell survival and viral replication. *Journal of Virology* **79**: 8742-8749.
- **Manns M.P., Foster G.R., Rockstroh J.K., Zeuzem S., Zoulim F., Houghton M. (2007).** The way forward in HCV treatment—finding the right path. *Nat. Rev. Drug Discov.* **6**(12):991-1000.
- **Maree A.F., Keulen W., Boucher C.A., De Boer R.J. (2000).** Estimating relative fitness in viral competition experiments. *J. Virol.* **74**: 11067–11072.
- **Markov P.V., Pepin J., Frost E., Deslandes S., Labbé A.C., Pybus O.G. (2009).** Phylogeography and molecular epidemiology of hepatitis C virus genotype 2 in Africa. *J. Gen. Virol.* **90**: 2086-2096.
- **Martell M., Esteban J.I., Quer J., Genescà J., Weiner A., Esteban R., Guardia J., Gómez J. (1992).** Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *Journal of Virology* **66**: 3225-3229.
- **Matsumoto M., Hwang S.B., Jeng K.S., Zhu N., Lai M.M. (1996).** Homotypic interaction and multimerization of hepatitis C virus core protein. *Virology* **218**:43-51.
- **McCaffrey K, Gouklani H, Boo I, Pountourios P, Drummer H.E. (2011).** The variable regions of hepatitis C virus glycoprotein E2 have an essential structural role in glycoprotein assembly and virion infectivity. *J. Gen. Virol.* **92**: 112-21.
- **McHutchinson J.G., Gondon S.C., Schiff E.R., et al. (1998).** Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis. *N. Engl. J. Med.* **339**: 1485–1492.
- **Méndez-Sánchez N., Gutiérrez-Grobe Y., Kobashi-Margáin R.A. (2010).** Epidemiology of HCV infection in Latin America. *Ann. Hepatol.* **9**: 27-29.
- **Meunier J.C., Russell R.S., Goossens V., Priem S., Walter H., Depla E., Union A., Faulk K. N., Bukh J., et al. (2008).** Isolation and characterization of broadly

neutralizing human monoclonal antibodies to the e1 glycoprotein of hepatitis C virus. *J. Virol.* **82**: 966–973.

- **Mindikoglu A.L., Miller R.R. (2009).** Hepatitis C in the elderly: epidemiology, natural history, and treatment. *Clin. Gastroenterol. Hepatol.* **7**:128-134.
- **Moradpour D., Penin F., Rice C.M. (2007).** Replication of hepatitis C virus. *Nature Review Microbiology* **5**: 453– 463
- **Moreau I., Hegarty S., Levis J., Sheehy P., Crosbie O., Kenny-Walsh E., Fanning L. (2006).** Serendipitous identification of natural Intergenotypic recombinants of hepatitis C in Ireland. *Virol. J.* **3**: 95.
- **Morel V., Descamps V., Frantois C., Fournier C., Brochot E., Capron D., Duverlie G., Castelain S. (2010).** Emergence of a genomic variant of the recombinant 2k/1b strain during a mixed Hepatitis C infection: A case report. *J. Clin. Virol.* **47**: 382–386.
- **Moreno M.P., Casane D., Lopez L., Cristina J. (2006).** Evidence of recombination in quasispecies populations of an Hepatitis C virus patient undergoing anti-viral therapy. *Virol. J.* **3**: 87.
- **Morikawa K., Lange C.M., Gouttenoire J., Meylan E., Brass V., Penin F., Moradpour D. (2011).** Nonstructural protein 3-4A: the Swiss army knife of hepatitis C virus. *J. Viral. Hepat.* **18(5)**: 305-315.
- **Moriya K., Fujie H., Shintani Y., Yotsuyanagi H., Tsutsumi T., Ishibashi K., Matsuura Y., Kimura S., Miyamura T., Koike K. (1998).** The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nat. Med.* **4**: 1065-1067.
- **Morrison D. (2005).** Networks in phylogenetic analysis: new tools for population biology. *Int. J. Parasitol.* **35**: 567–582.
- **Muñoz E., Park J.M., Deem M.W. (2008).** Quasispecies theory for horizontal gene transfer and recombination. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **78**: 061921.
- **Murphy M.D., Rosen H.R., Marousek G.I., Chou S. (2002).** Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 regions as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. *Dig. Dis. Sci.* **47**: 1195–1205.

- **Netski D., Mosbrugger T.J., Astemborski J., et al. (2007).** CD4+ T cell-dependent reduction in hepatitis C virus-specific humoral immune responses after HIV infection. *J. Infect. Dis.* **195**(6): 857–863.
- **Newman M.E.J. (2002).** Assortative Mixing in Networks. *Physical Review Letters* **89** (20): 208701.
- **Nousbaum J.B., Polyak S.J., C. Ray S., Sullivan D.G., Larson A.M., Carithers R.L., Gretch D.R., (2000).** Prospective Characterization of Full-Length Hepatitis C Virus NS5A Quasispecies during Induction and Combination Antiviral Therapy. *J. Virology* **74**: 9028–9038.
- **Nowak M.A. (2006).** Evolutionary dynamics. Belknap Press, Cambridge, MA.
- **Offman M.N., Tournier A.L., Bates P.A. (2008).** Alternating evolutionary pressure in a genetic algorithm facilitates protein model selection BMC Structural Biology, **8**: 34.
- **Ojosnegros S. et al. (2011).** Viral genome segmentation can result from a trade-off between genetic content and particle stability. *PLoS Genet.* **7**:e1001344.
- **Olzin T., Daneschmand S.V. (2003).** On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* **31**: 12-20.
- **Orland J.R., Wright T.L., Cooper S. (2001).** Acute hepatitis C. *Hepatology* **33**: 321-327.
- **Orlent H., Desombere I., Hansen B., Van Vlierberghe H., Haagmans B., De Knecht R.J., Schalm S.W., Leroux-Roels G., Janssen H.L. DITTO-HCV Study Group. (2010).** Baseline anti-NS4a antibodies in combination with on-treatment quantitative HCV-RNA reliably identifies nonresponders to pegylated interferon-ribavirin combination therapy after 4 weeks of treatment. *Eur. J. Gastroenterol. Hepatol.* **22**: 1443-1448.
- **Ospina-Bedoya M., Campillo-Pedroza N., Franco-Salazar J.P., Gallego-Gómez J.C. (2014).** Computational identification of Dengue virus microRNA-like structures and their cellular targets. *Bioinf. Biol. Insights.* **8**:169–176.
- **Owsianka A., Clayton R.F., Loomis-Price L.D., McKeating J.A., Patel A.H. (2001).** Functional analysis of hepatitis C virus E2 glycoproteins and virus-like particles reveals structural dissimilarities between different forms of E2. *J. Gen. Virol.* **82**: 1877–1883.

- **Owsianka A., Tarr A. W., Juttla V.S., Lavillette D., Bartosch B., Cosset F.L., Ball, J.K., Patel A.H. (2005).** Monoclonal antibody AP33 defines a broadly neutralizing epitope on the hepatitis C virus E2 envelope glycoprotein. *J. Virol.* **79**: 11095–11104.
- **Owsianka A.M., Timms J.M., Tarr A.W., Brown R. J. P., Hickling T. P., Szwejk A., Bienkowska-Szewczyk K., Thomson B.J., Patel A.H., Ball J.K. (2006).** Identification of conserved residues in the E2 envelope glycoprotein of the hepatitis C virus that are critical for CD81 binding. *J. Virol.* **80**: 8695–8704.
- **Park J., Kang W., Ryu S.W., et al. (2012).** Hepatitis C virus infection enhances TNF α -induced cell death via suppression of NF- κ B. *Hepatology* **56**: 831-840.
- **Pawlotsky J.M. (2003).** The nature of interferon-alfa resistance in hepatitis C virus infection. *Curr. Opin. Infect. Dis.* **16**: 587–592.
- **Pawlotsky J.M. (2006)** Hepatitis C virus population dynamics during infection. *Current Topical Microbiology Immunology* **299**: 261–284.
- **Penin F., Dubuisson J., Rey F., Moradpour D., Pawlotsky J., (2004).** Structural Biology of Hepatitis C Virus. *Hepatology* **39**: 5-19.
- **Perales C., Lorenzo-Redondo R., López-Galíndez C., Martínez M.A., Domingo E. (2010).** Mutant spectra in virus behavior. *Future Virol.* **5**: 679–698.
- **Perales C., Martin V., Ruiz-Jarabo C.M., Domingo E. (2005).** Monitoring sequence space as a test for the target of selection in viruses. *J. Mol. Biol.* **345**: 451–459.
- **Perotti M., Mancini N., Diotti R. A., Tarr A. W., Ball J. K., Owsianka A., Adair R., Patel A.H., Clementi M., Burioni, R. (2008).** Identification of a broadly cross-reacting and neutralizing human monoclonal antibody directed against the hepatitis C virus E2 protein. *J. Virol.* **82**: 1047–1052.
- **Perz J.F., Armstrong G.L., Farrington L.A., Hutin Y.J., Bell B.P. (2006).** The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J. Hepatol.* **45(4)**: 529-538.
- **Petta S., Craxì A. (2011).** Optimal therapy in hepatitis C virus genotypes 2 and 3 patients. *Liver Int.* **1**: 36-44.
- **Pfeffer S., Baumert T.F. (2010).** Impact of microRNAs for pathogenesis and treatment of hepatitis C virus infection. *Gastroenterol Clin. Biol.* **34**: 431–435.

- **Pfeffer S., Zavolan M., Grässer F.A., Chien M., Russo J.J., Ju J., John B., Enright A.J., Marks D., Sander C., Tuschl T. (2004).** Identification of virus-encoded microRNAs. *Science* **304**:734–736.
- **Pileri P., Uematsu Y., Campagnoli S., Galli G., Falugi F., Petracca R., Weiner A.J., Houghton M., Rosa D., Grandi G., Abrignani S. (1998).** Binding of hepatitis C virus to CD81. *Science* **282**: 938-941.
- **Plakunmonthon S., Panjaworayan N., Thienprasert T., Khongnomnan K., Poovorawan Y., Payungporn S. (2014).** Computational prediction of hybridization patterns between hepatitisC viral genome and human microRNAs. *J. Comput. Sci.* **5**: 327.
- **Polyak S.J., Khabar K.S.A., Paschal D.M., Ezelle H.J., Duverlie G., Barber G.N., Levy D.E., Mukaida N., Gretch D.R. (2001).** Hepatitis C virus nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response. *J. Virol.* **75**: 6095–6106.
- **Poon A.F.Y., Lewis F., Frost S.D.W., Kosakovsky-Pond S.L. (2008).** Spidermonkey: rapid detection of co-evolving sites using Bayesian graphical models. *Bioinformatics* **24**: 1949-1950.
- **Poon A.F.Y., Lewis F.I., Kosakovsky-Pond S.L., Frost S.D.W. (2007).** An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. *PLoS Comput. Biol.* **3**: e231.
- **Posada D. and Crandall K. (2001).** Intraspecific gene genealogies: trees grafting into networks. *Trends. Ecol. Evol.* **16**: 37–45.
- **Poss M., Rodrigo A.G., Gosink J.J., Learn G.H., de Vange P.D., Martin H.L. Jr, Bwayo J., Kreiss J.K., Overbaugh J. (1998).** Evolution of envelope sequences from the genital tract and peripheral blood of women infected with clade A human immunodeficiency virus type 1. *J. Virol.* **72**: 8240–8251.
- **Poynard T., Marcellin P., Lee S.S., Niederau C., Minuk G.S., Ideo G., Bain V., Heathcote J., Zeuzem S., Trepo C., Albrecht J. (1998).** Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. *Lancet* **352**: 1426–1432.
- **Prentoe J., Jensen T. B., Meuleman P., Serre S.B. N., Scheel T.K.H., Leroux-Roels G., Gottwein J.M., Bukh J. (2011).** Hypervariable region 1 differentially

impacts viability of hepatitis C virus strains of genotypes 1 to 6 and impairs virus neutralization. *J. Virol.* **85**: 2224–2234.

- **Puig-Basagoiti F., Forns X., Furci I., Ampurdanes S., Gimenez-Barcons M., Franco S., Sanchez-Tapias J.M., Saiz J.C. (2005).** Dynamics of hepatitis C virus NS5A quasispecies during interferon and ribavirin therapy in responder and non-responder patients with genotype 1b chronic hepatitis C. *J. Gen. Virol.* **86**: 1067–1075.
- **Puig-Basagoiti F., Saiz J.C., Forns X., et al. (2001).** Influence of the genetic heterogeneity of the ISDR and PePHD regions of hepatitis C virus on the response to interferon therapy in chronic hepatitis C. *J. Med. Virol.* **65**: 35–44.
- **Quiñones-Mateu M.E., Arts E. (2006).** Virus fitness: concept, quantification, and application to HIV population dynamics. *Curr. Top. Microbiol. Immunol.* **299**: 83–140.
- **Ramachandran S., Campo D.S., Dimitrova Z.E., Xia G.L., Purdy M.A., Khudyakov Y.E. (2011).** Temporal variations in the hepatitis C virus intrahost population during chronic infection. *Journal of Virology*, **85**: 6369–6380.
- **Ray S.C., Fanning L., Wang X.H., Netski D.M., Kenny-Walsh E., Thomas D.L. (2005).** Divergent and convergent evolution after a common-source outbreak of hepatitis C virus. *J. Exp. Med.* **201**: 1753–1759.
- **Ray S.C., Wang Y.M., Laeyendecker O., Ticehurst J.R., Villano S.A., Thomas D.L. (1999).** Acute hepatitis C virus structural gene sequences as predictors of persistent viremia: hypervariable region 1 as a decoy. *J. Virol.* **73**: 2938–2946.
- **Reed K.E., Grakoui A., Rice C.M. (1995).** Hepatitis C virus encoded NS2-3 protease: cleavage-site mutagenesis and requirements for biomolecular cleavage. *J. Virol.* **69**: 4127–4136.
- **Reyes G.R. (2002).** The nonstructural NS5A protein of hepatitis C virus: an expanding, multifunctional role in enhancing hepatitis C virus pathogenesis. *J. Biomed. Sci.* **9**: 187–197.
- **Roque Afonso A.M., Jiang J., Penin F., Tareau C., Samuel D., Petit M.A., Bismuth H., Dussaix E., Feray C. (1999).** Nonrandom distribution of hepatitis C virus quasispecies in plasma and peripheral blood mononuclear cell subsets. *J. Virol.* **73**: 9213–9221.

- **Rosen H.R., Gretch D.R. (1999).** Hepatitis C virus: current understanding and prospects for future therapies. *Mol. Med. Today.* **9:** 393-399.
- **Rosewick N., Momont M., Durkin K., Takeda H., Caiment F., Cleuter Y., et al. (2013).** Deep sequencing reveals abundant noncanonical retroviral microRNAs in B-cell leukemia/lymphoma. *Proc. Natl. Acad. Sci. U S A* **110:**2306–2311.
- **Rothman A.L., Morishima C., Bonkovsky H.L., et al. (2005).** HALTC Trial Group. Associations among clinical, immunological, and viral quasispecies measurements in advanced chronic hepatitis C. *Hepatology* **41(3):** 617–625.
- **Ruiz-Jarabo C.M., Arias A., Baranowski E., Escarmís C., Domingo E. (2000).** Memory in viral quasispecies. *J. Virol.* **74:** 3543–3547.
- **Scarselli E., Ansuini H., Cerino R., Roccasecca R.M., Acali S., Filocamo G., Traboni C., Nicosia A., Cortese R., Vitelli A. (2002).** The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* **21:** 5017–5025.
- **Schinkel J., Spoon W.J., Kroes A.C. (2004).** Meta-analysis of mutations in the NS5A gene and hepatitis C virus resistance to interferon therapy: uniting discordant conclusions. *Antivir. Ther.* **9:** 275–286.
- **Schuster C., Isel C., Imbert I., et al. (2002).** Secondary Structure of the 3' Terminus of Hepatitis C Virus Minus-Strand RNA. *Journal of Virology* **76:** 8058-8068.
- **Sentandreu V., Jiménez-Hernández N., Torres-Puente M., Bracho M.A., Valero A., Gosalbes M.J., Ortega E., Moya A., González-Candelas F. (2008).** Evidence of recombination in inpatient populations of Hepatitis C Virus. *PLoS One.* **3:** e3239.
- **Seo G.J., Chen C.J., Sullivan C.S: (2009).** Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. *Virology* **383:**183–187.
- **Severson W.E., Schmaljohn C.S., Javadian A., Jonsson C.B. (2003).** Ribavirin causes error catastrophe during Hantaan virus replication. *J. Virol.* **77:** 481–488.
- **Sharma S.A., Feld J.J. (2014).** Acute hepatitis C: management in the rapidly evolving world of HCV. *Curr. Gastroenterol. Rep.* **16(2):** 371.
- **Sharma S.D. (2010).** “Hepatitis c virus: molecular biology & current therapeutic options,” *Indian Journal of Medical Research*, vol. 131, n° 1, pp. 17–34.

- **Shepard C.W., Finelli L., Alter M.J. (2005).** Global epidemiology of hepatitis C virus infection. *Lancet. Infect. Dis.* **5**: 558-567.
- **Shi J., Duan Z., Sun J., Wu M., Wang B., Zhang J., Wang H., Hu N., Hu Y. (2014).** Identification and validation of a novel microRNA-like molecule derived from a cytoplasmic RNA virus antigenome by bioinformatics and experimental approaches. *J.Virol.* **11**:121.
- **Shi J., Duan Z., Sun J., Wu M., Wang B., Zhang J., Wang H., Hu N., Hu Y. (2014).** Identification and validation of a novel microRNA-like molecule derived from a cytoplasmic RNA virus antigenome by bioinformatics and experimental approaches. *Proceedings of the National Academy of Science U S A*, **110**: 228-233.
- **Shimakami T., Hijikata M., Luo H., Ma Y.Y., Kaneko, S., et al. (2004).** Effect of interaction between hepatitis C virus NS5A and NS5B on hepatitis C virus RNA replication with the hepatitis C virus replicon. *J. Virol.* **78**: 2738-2748.
- **Shimizu Y.K., Igarashi H., Kiyohara T., Cabezon T., Farci P., Purcell R.H., Yoshikura H. (1996).** A hyperimmune serum against a synthetic peptide corresponding to the hypervariable region 1 of hepatitis C virus can prevent viral infection in cell cultures. *Virology* **223**: 409–412.
- **Shiratori Y., Shiina S., Teratani T., Imamura M., Obi S., Sato S., Koike Y., Yoshida H., Omata M. (2003).** Interferon Therapy after Tumor Ablation Improves Prognosis in Patients with Hepatocellular Carcinoma Associated with Hepatitis C Virus. *Annals of Internal Medicine* **138**: 299–306.
- **Shrivastava S., Mukherjee A., Ray R.B. (2013).** Hepatitis C virus infection, microRNA and liver disease progression. *World Journal of Hepatology* **5**:479-486.
- **Shrivastava S., Steele R., Ray R., Ray R.B. (2015).** MicroRNAs: Role in hepatitis C virus pathogenesis. *Genes Dis.* 2015.
- **Simmonds P. (2004).** Genetic diversity and evolution of hepatitis C virus-15 years on. *J. Gen. Virol.* **85**: 3173–3188.
- **Simmonds P., Bukh J., Combet C., Deleage G., Enomoto N., Feinstone S., Halfon P., Inchauspe G., Kuiken C., Maertens G., Mizokami M., Murphy D.G., Okamoto H., Pawlotsky J.M., Penin F., Sablon E., Shin I.T., Stuyver L.J., Thiel H.J., Viazov, S., Weiner A.J., Widell A. (2005).** Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hematology* **42**: 962–973.

- **Smith D.B., Bukh J., Kuiken C., Muerhoff A.S., Rice C.M., Stapleton J.T., Simmonds P. (2014).** Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment *Web resource. Hepatology*. **59**: 318-327.
- **Smith R. M., Walton C. M., Wu C. H. et al. (2002).** Secondary structure and hybridization accessibility of Hepatitis C Virus 3' terminal sequences. *Journal of Virology* **76**: 9563–9574.
- **Stadler, T. (2010).** Sampling-Through-Time in Birth-Death Trees. *Journal of Theoretical Biology*, **267**, 396-404.
- **Stadler, T. (2011).** Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Science USA*, **108**: 6187-6192.
- **Stadler, T., Kouyos, R., von Wyl, V., Yerly, S., Boni, J., Burgisser, P., et al. (2012).** Estimating the Basic Reproductive Number from Viral Sequence Data. *Molecular Biology and Evolution*, **29**: 347-357.
- **Stadler, T., Kühnert, D., Bonhoeffer, S. and Drummond, A.J. (2013).** Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proceedings of the National Academy of Science USA*, **110**: 228-233.
- **Steel M.A. (2005).** Should phylogenetic models be trying to fit an elephant? *Trends. Genet.* **21**: 307–309.
- **Sullivan C.S., Ganem D. (2005).** MicroRNAs and viral infection. *Mol. Cell.* **1**: 3-7.
- **Sy T., Jamal M.M. (2006).** Epidemiology of hepatitis C virus (HCV) infection. *Int. J. Med. Sci.* **3**: 41-46.
- **Szabo G. (2006).** Hepatitis C virus NS5A protein, a master regulator? *Gastroenterology*. **130**: 995–998.
- **Szabo S.M., Bibby M., Yuan Y., Donato B.M., Jiménez-Mendez R., Castañeda-Hernández G., Rodríguez-Torres M., Levy A.R. (2012).** The epidemiologic burden of hepatitis C virus infection in Latin America. *Ann. Hepatol.* **11**: 623-635.
- **Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.. (1991).** Structure and organization of the hepatitis C virus genome isolated from human carriers. *J. Virol.* **65(3)**: 1105-1113.

- **Takikawa S., Ishii K., Aizaki H., Suzuki T., Asakura H., Matsuura Y., Miyamura T., (2000).** Cell fusion activity of Hepatitis C virus envelope proteins. *J. Virol.* **65**: 1105-1113.
- **Tam R.C., Lau J.Y., Hong Z., 2001.** Mechanisms of action of ribavirin in antiviral therapies. *Antivir. Chem. Chemother.* **12**: 261–272.
- **Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. (2011)** MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**: 2731-2739.
- **Tan S.L., Katze M.G. (2001).** How hepatitis C virus counteracts the interferon response: the jury is still out on NS5A. *Virology* **284**: 112.
- **Thibault P.A., Wilson J.A. (2013).** Targeting miRNAs to treat hepatitis C virus infections and liver pathology: inhibiting the virus and altering the host. *Pharm. Res.* **75**: 48–59.
- **Thirugnanasambantham K., Hairul-Islam V.I., Saravanan S., Subasri S., Subastri A. (2013).** Computational approach for identification of *Anopheles gambiae* miRNA involved in modulation of host immune response. *Appl. Biochem. Biotechnol.* **170**:281–291..
- **Thomas E., Gonzalez V.D., Qisheng L., et al. (2011).** HCV Infection Induces a Unique Hepatic Innate Immune Response Associated With Robust Production of Type III Interferons. *Gastroenterology.* **142**: 978–988.
- **Timpe J.M., Stamataki Z., Jennings A., Hu K., Farquhar M.J., Harris H.J., Schwarz A., Desombere I., Roels G.L., et al. (2008).** Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology* **47**: 17–24.
- **Tsukiyama-Kohara K., Iizuka N., Kohara M., Nomoto A. (1992).** Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* **66**: 1476-1483.
- **Van Doorn L.J., Capriles I., Maertens G., De Leys R., Murray K., Kos T., Schellekens H., Quint W. (1995).** Sequence evolution of the hypervariable region in the putative envelope region E2/NS1 of hepatitis C virus is correlated with specific humoral immune responses. *Journal of Virology*, **69**: 773–778.
- **Verstrepen B.E., Depla E., Rollier C.S., Mares G., Drexhage J.A.R., Priem S., Verschoor E.J., Koopman G., Granier C., et al. (2011).** Clearance of genotype 1b

hepatitis C virus in chimpanzees in the presence of vaccine-induced E1-neutralizing antibodies. *J. Infect. Dis.* **204** : 837–844.

- **Vignuzzi M., Stone J.K., Andino R. (2005).** Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Res.* **107**: 173–181.
- **Voisset C., Callens N., Blanchard E., Op De Beeck A., Dubuisson J., Vu-Dac N. (2005).** High density lipoproteins facilitate hepatitis C virus entry through the scavenger receptor class B type I. *J. Biol. Chem.* **280**: 7793–7799.
- **Walewski J.L., Keller T.R., Stump D.D., Branch A.D. (2001).** Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA.* **7(5)**: 710-721.
- **Wang K.C., Chang H.Y. (2011).** Molecular mechanisms of long noncoding RNAs. *Mol. Cell.* **43**: 904–914.
- **Wang W., Lahser F.C., Yi M., Wright-Minogue J., Xia E., Weber P.C., Lemon S.M., Malcolm B.A. (2004).** Conserved C-Terminal-Threonine of Hepatitis C virus NS3 regulate autoproteolysis and prevents product inhibition. *J. Virol.* **78**: 700-709.
- **Weiner A.J., Thaler M.M., Crawford K., Ching K., Kansopon J., Chien D.Y., Hall J.E., Hu F., Houghton M. (1993).** A unique, predominant hepatitis C virus variant found in an infant born to a mother with multiple variants. *J. Virol.* **67**: 4365–4368.
- **WHO. (2003)** Hepatitis C.
- **Wimmer E., Paul A.V. (2010).** The making of a picornavirus genome, p 33–58. In Ehrenfeld E, Domingo E, Roos RP (ed), *The picornaviruses*. ASM Press, Washington, DC.
- **Witherell G.W., Beineke P. (2001).** Statistical analysis of combined substitutions in non-structural 5A region of hepatitis C virus and interferon response. *J. Med. Virol.* **63**: 8–16.
- **Witteveldt J., Evans M.J., Bitzegeio J., Koutsoudakis G., Owsianka A.M., Angus A.G.N., Keck Z.Y., Fong S.K.H., Pietschmann T., et al. (2009).** CD81 is dispensable for hepatitis C virus cell-to-cell transmission in hepatoma cells. *J. Gen. Virol.* **90**: 48–58.

- **World Health Organization.** Initiative for vaccine research. Hepatitis C. Available online: http://www.who.int/vaccine_research/diseases/viral_cancers/en/index2.html (accessed on 12 September 2011).
- **Wright F., (1990).** The “effective number of codons” used in a gene. *Gene* **87**:23-29.
- **Yang D.R., Zhu H.Z. (2015).** Hepatitis C virus and antiviral innate immunity: Who wins at tug-of-war?. *World Journal of Gastroenterology* **21**: 3786-3800.
- **Yeo A.E., Ghany M., Conry-Cantilena C., et al. (2001).** Stability of HCV-RNA level and its lack of correlation with disease severity in asymptomatic chronic hepatitis C virus carriers. *J. Viral Hepat.* **8(4)**:256–263.
- **Young K.C., Lindsay K.L., Lee K.J., Liu W.C., He J.W., Milstein S.L., Lai M.M. (2003).** Identification of a ribavirin-resistant NS5B mutation of hepatitis C virus during ribavirin monotherapy. *Hepatology* **38**: 869–878.
- **Yuki N., Hayashi N., Moribe T., et al. 1997.** Relation of disease activity during chronic hepatitis C infection to complexity of hypervariable region 1 quasispecies. *Hepatology* **25(2)**: 439–444.
- **Zagordi O., Bhattacharya A., Eriksson N., Beerenwinkel N. (2011).** ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. *BMC Bioinformatics* **12**: 119.
- **Zaltron S., Spinetti A., Biasi L., Baiguera C., Castelli F. (2012).** Chronic HCV infection: epidemiological and clinical relevance. *BMC Infect. Dis.* **12 2**: S2.
- **Zeng Q.L., Zhang J.Y., Zhang Z., Wang L.F., Wang F.S. (2013).** Sofosbuvir and ABT-450: terminator of hepatitis C virus? *World J. Gastroenterol.* **19(21)**: 3199-3206.
- **Zeuzem S. (2004).** Heterogenous virologic response rates to interferon-based therapy in patients with chronic hepatitis C: who responds less well? *Ann. Intern. Med.* **140**: 370-381.
- **Zhang Y., Fan M., Geng G., Liu B., Huang Z., Luo H., et al. (2014).** A novel HIV-1-encoded microRNA enhances its viral replication by targeting the TATA box region. *Retrovirology.* **11**:23.
- **Zhang Y., Jamaluddin M., Wang S., Tian B., Garofalo R.P., Casola A., Brasier A.R. (2003).** Ribavirin treatment up-regulates antiviral gene expression via the

interferon-stimulated response element in respiratory syncytial virus-infected epithelial cells. *J. Virol.* **77**: 5933–5947.