



Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

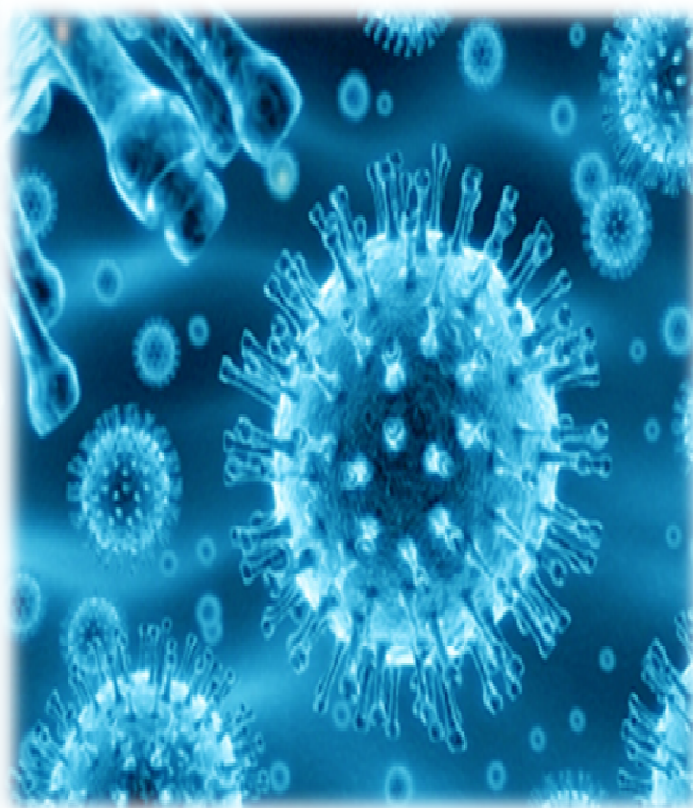
ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. The access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

**ESTUDIO DE
COMPLEJIDAD Y
EVOLUCIÓN DEL GEN NS3
DEL VIRUS DE LA
HEPATITIS C (VHC)
MEDIANTE
ULTRA-DEEP
PYROSEQUENCING
(UDPS) EN PACIENTES
EXPUESTOS AL MISMO
INÓCULO VIRAL**

M^a DOLORES CUBERO LEÓN

BARCELONA 2012





Universitat Autònoma de Barcelona
Departament de Bioquímica i Biologia Molecular

**ESTUDIO DE COMPLEJIDAD Y EVOLUCIÓN DEL GEN NS3 DEL VIRUS
DE LA HEPATITIS C (VHC) MEDIANTE ULTRA-DEEP
PYROSEQUENCING (UDPS) EN PACIENTES EXPUESTOS AL MISMO
INÓCULO VIRAL**

Memoria presentada por María Dolores Cubero León para optar al grado de Doctor por la
Universitat Autònoma de Barcelona

Directores de Tesis

Josep Quer Sivila
Juan Ignacio Esteban

Tutor de Tesis

Jaume Farrés Vicén

Laboratori Malalties Hepàtiques, Vall d'Hebron Institut de Recerca (VHIR)

Barcelona 2012



Universitat Autònoma de Barcelona

Departament de Bioquímica i Biologia Molecular



**ESTUDIO DE COMPLEJIDAD Y EVOLUCIÓN DEL GEN NS3 DEL VIRUS
DE LA HEPATITIS C (VHC) MEDIANTE ULTRA-DEEP
PYROSEQUENCING (UDPS) EN PACIENTES EXPUESTOS AL MISMO
INÓCULO VIRAL**

Memoria presentada por María Dolores Cubero León para optar al grado de Doctor por la
Universitat Autònoma de Barcelona

Directores de Tesis

Josep Quer Sivila

Juan Ignacio Esteban

Estudio financiado por la beca SAF2009-10403 del MINECO, Marató TV3 052310, y por
CIBERehd

Laboratori Malalties Hepàtiques, Vall d'Hebron Institut de Recerca (VHIR)

Barcelona 2012



ÍNDICE

1	<u>ABREVIATURAS</u>	1
2	<u>INTRODUCCIÓN</u>	3
2.1	GENOMA VHC	3
2.2	VARIABILIDAD GENÉTICA DEL VHC. GENOTIPOS Y SUBTIPOS	5
2.3	DINÁMICA DE LOS VIRUS DE ARN. QUASIESPECIES	6
2.4	INFECCIÓN POR VHC	10
2.5	FACTORES VIRALES QUE INFLUYEN EN LA EVOLUCIÓN DE LA INFECCIÓN POR VHC HACIA CRONICIDAD O CURACIÓN	11
2.5.1	COMPLEJIDAD DE LA POBLACIÓN VIRAL	12
2.5.2	MODO DE TRANSMISIÓN DEL VHC	12
2.5.3	EVASIÓN DEL SISTEMA INMUNE	13
2.5.4	RESISTENCIA A DROGAS ANTIVIRALES	14
2.6	FACTORES DEL PACIENTE QUE INFLUYEN EN LA EVOLUCIÓN DE LA INFECCIÓN POR VHC HACIA CRONICIDAD O CURACIÓN	15
2.6.1	IL-28B	16
2.6.2	SISTEMA INMUNE	16
2.6.3	OTROS FACTORES	17
2.7	PLATAFORMAS DE SECUENCIACIÓN “HIGH-THROUGHPUT”	18
3	<u>ESTUDIO 1:”EFFECT OF BOTTLENECKING ON EVOLUTION OF THE NONSTRUCTURAL PROTEIN 3 GENE OF HEPATITIS C VIRUS DURING SEXUALLY TRANSMITTED ACUTE RESOLVING INFECTION”</u>	22
3.1	HIPÓTESIS:	22
3.2	OBJETIVO:	22
3.3	RESUMEN:	22
4	<u>ESTUDIO 2:”NATURALLY OCCURRING NS3-PROTEASE-INHIBITOR RESISTANT MUTANT A156T IN THE LIVER OF AN UNTREATED CHRONIC HEPATITIS C PATIENT”</u>	36
4.1	HIPÓTESIS	36
4.2	OBJETIVO	36
4.3	RESUMEN	36
5	<u>ESTUDIO 3: IDENTIFICACIÓN DE FACTORES VIRALES Y DEL PACIENTE INVOLUCRADOS EN RESOLUCIÓN O CRONICIDAD</u>	48
5.1	INTRODUCCIÓN	48
5.1.1	FACTORES VIRALES (HUÉSPED)	48
5.1.2	FACTORES GENÓMICOS DEL PACIENTE (HOSPEDADOR).	48
5.2	OBJETIVOS	48
5.3	RESULTADOS	49
5.3.1	HAPLOTIPOS HLA	49
5.3.2	GENOTIPO IL28B	50

5.3.3	ANÁLISIS A NIVEL DE NUCLEÓTIDOS DE LA COMPLEJIDAD VIRAL DE LA QUASIESPECIE:	51
5.3.4	ANÁLISIS A NIVEL DE AMINOÁCIDOS DE LA COMPLEJIDAD VIRAL DE LA QUASIESPECIE:	55
5.3.5	DN/DS POBLACIONAL	59
5.3.6	INTERACCIÓN ENTRE FACTORES VIRALES Y DEL PACIENTE. VARIABILIDAD Y RESPUESTA INMUNE ADAPTATIVA CD4 ESPECÍFICA	60
5.3.7	DENSIDAD DE MUTACIÓN POBLACIONAL (MONTSERRAT PLOTS)	63
5.3.8	IDENTIDAD POBLACIONAL MEDIANTE COEFICIENTE DE BHATTACHARYYA	68
5.3.9	DISTANCIA GENÉTICA	70
5.3.10	ÁRBOLES FILOGENÉTICOS	71
5.3.11	COMPARACIÓN ANÁLISIS FILOGENÉTICOS Y ESTRUCTURA DE LA DENSIDAD DE MUTACIÓN DE LOS PACIENTES A Y B	95
5.3.12	ANÁLISIS DE MUTACIONES DE RESISTENCIA CONTRA INHIBIDORES DE NS3 PROTEASA DEL VHC	103
6	<u>DISCUSIÓN</u>	105
6.1	FACTORES DEL PACIENTE (IL28 Y HLA)	105
6.2	FACTORES VIRALES	106
7	<u>CONCLUSIONES</u>	114
8	<u>MATERIALES Y MÉTODOS</u>	116
8.1	PACIENTES	116
8.2	TÉCNICAS DE BIOLOGÍA MOLECULAR 1^{ER} Y 2^º ESTUDIO	120
8.2.1	HOMOGENEIZACIÓN DE LA MUESTRA	120
8.2.2	CONTROL DE CANTIDAD Y CALIDAD DE ARN EXTRAÍDO	121
8.2.3	RT-PCR-NESTED	122
8.3	TÉCNICAS DE BIOLOGÍA MOLECULAR 3^{ER} ESTUDIO	129
8.3.1	IL28 Y HLA	129
8.3.2	ANTÍGENOS DEL VHC	129
8.3.3	ENSAYO DE ENZYME-LINKED IMMUNOSPOT (ELISPOT)	130
8.3.4	ANÁLISIS DE LA REGIÓN NS3 MEDIANTE SECUENCIACIÓN MASIVA	131
8.3.5	ALGORITMO PARA INTERPRETAR LAS SECUENCIAS	136
8.3.6	ANÁLISIS DE LAS SECUENCIAS	137
9	<u>BIBLIOGRAFÍA</u>	149

1 ABREVIATURAS

AA, Aminoácido

ARN, Ácido ribonucleico

ADN, Ácido desoxirribonucleico

ALT, Alanina aminotransferasa

BOC, Boceprevir

CMH, Complejo mayor de histocompatibilidad

DAAs, Direct antiviral agents

ELISPOT, Enzyme-Linked Immunosorbent spot

EMA, European Medicines Agency

emPCR, Emulsion polymerase chain reaction

GS-FLX, Genome Sequencer-FLX

IFN, Interferón

HLA, Human leukocyte antigen

NT, Nucleótido

ORF, Open Reading Frame

PCR, Polymerase chain reaction

PegIFN, Interferón pegilado

RBV, Ribavirina

RdRp-RNA dependent RNA polymerase

SOC, Standar of care

SNP, Single nucleotide polymorphism

RVS, Respuesta virológica sostenida

TAC, Tomografía axial computarizada

Th1, Respuesta inmunológica tipo helper 1

TVR, Telaprevir

UDPS, Ultra Deep Pyrosequencing

UPGMA, Unweighted pair-group method with arithmetic means

UTR, Untranslated región

VHC, Virus de la hepatitis C

WT, Wild type

INTRODUCCIÓN

2 INTRODUCCIÓN

2.1 Genoma VHC

El virus de la hepatitis C (VHC) es un virus de ARN de polaridad positiva, lineal, de cadena sencilla, perteneciente al género *Hepacivirus* de la familia *Flaviviridae*. El VHC tiene un tamaño de 9.5 kb con una sola pauta de lectura abierta (open reading frame- ORF-) que se traduce en una poliproteína de aproximadamente 3.011 aminoácidos (aa), en cuyos extremos se encuentran dos regiones no-codificantes (5'-UTR y 3'-UTR). La traducción del genoma del VHC resulta en un precursor poliproteico codificado por el ORF que es procesado co- y post-traduccionalmente en el retículo endoplasmático (1) tanto por proteasas celulares como virales (Figura 1) para producir 10 proteínas maduras C-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B (2-4). El tercio N-terminal de la poliproteína da lugar a las proteínas estructurales Core, E1 y E2 que forman la partícula viral. Separando esta región estructural de la no-estructural, se encuentra la proteína p7 que es un polipéptido que pertenece a la familia de las viroporinas y se ubica en el retículo endoplasmático como una proteína de membrana integral y tiene una función crucial en el ensamblaje y liberación de los viriones infecciosos (5;6). El procesamiento de todas estas 4 proteínas (Core/E1, E1/E2, E2/p7 y p7/NS2) es llevado a cabo por peptidasas de la célula infectada. Los dos tercios C-terminales, codifican para las proteínas no estructurales esenciales para la replicación del ARN viral y son procesadas por dos proteasas virales. La proteína NS2 es una proteína transmembrana de 217 aminoácidos compuesta por un dominio N-terminal altamente hidrofóbico, un dominio C-terminal globular y un subdominio de proteasa citosólica, este último capaz de formar dímeros creando un sitio activo compuesto. Esta proteasa no es requerida directamente para la replicación del ARN, pero se autoescinde por la N-terminal de NS3 para la formación de una replicasa activa la NS3. NS2 recluta la glicoproteína de la envoltura en conjunción con p7 y probablemente LDL (Low density lipoprotein) y sirve como un organizador del ensamblaje de VHC mediante la

participación en múltiples interacciones proteína-proteína requeridas para la formación del virión (7). NS4B es una proteína integral de membrana de 261 aminoácidos localizada en el retículo endoplasmático, consta de 4 segmentos transmembrana en su parte central y está involucrada en el proceso de formación del complejo de replicación viral. NS4B induce una alteración específica de membrana, denominada “membranous web”, que sirve de soporte para el complejo de replicación (8). NS5A es una fosfoproteína anclada en el retículo endoplasmático esencial para la replicación y ensamblaje del ARN y, por tanto, en la producción de partículas virales se une a NS5B y activa a esta enzima e interactúa con múltiples proteínas celulares y virales. Está constituida por 3 dominios citoplasmáticos: D1, D2 y D3. D1 es un dominio de unión a zinc, que puede adoptar 2 tipos de estructuras: un dímero en forma de gancho adaptado para su unión al ARN, y una estructura alternativa probablemente requerida para otra función en la replicación, aunque todavía está por dilucidar la función completa de D1. D2 es esencial para la replicación del ARN y forma un punto de anclaje con la isomerasa cis/trans/prolil ciclofilina de la célula hospedadora, proporcionando una base molecular para el uso de inhibidores de la ciclofilina como estrategia antiviral (9). Finalmente D3 está relacionada con la producción y ensamblaje de partículas virales (10). La ARN polimerasa dependiente de ARN (RdRp) representada por la proteína no estructural 5B (NS5B) pertenece a una familia de proteínas de anclaje a la membrana. Forma un complejo de replicación con otras proteínas, como core, NS3, NS4A, NS4B y NS5A. El dominio C-terminal de la RdRp además de servir como punto de anclaje a la membrana, está implicado en interacciones intramembrana proteína-proteína. Por tanto, tiene implicaciones para la arquitectura funcional del complejo de replicación del virus, y además esta asociación con la membrana es esencial para la replicación viral (11). La RdRp reconoce señales de transcripción situadas en el extremo 3' no-codificante para dar lugar a la cadena negativa del ARN del virus, la cual servirá de molde para producir la cadena positiva que será incorporada a los viriones. NS3, además de actividad proteasa, tiene actividad NTPasa y helicasa, induciendo probablemente el despliegue de la molécula de ARN durante la replicación viral. Finalmente, la progenie viral es encapsidada en

asociación con membranas intracelulares y alcanza el medio extracelular a través del aparato secretor de la célula.

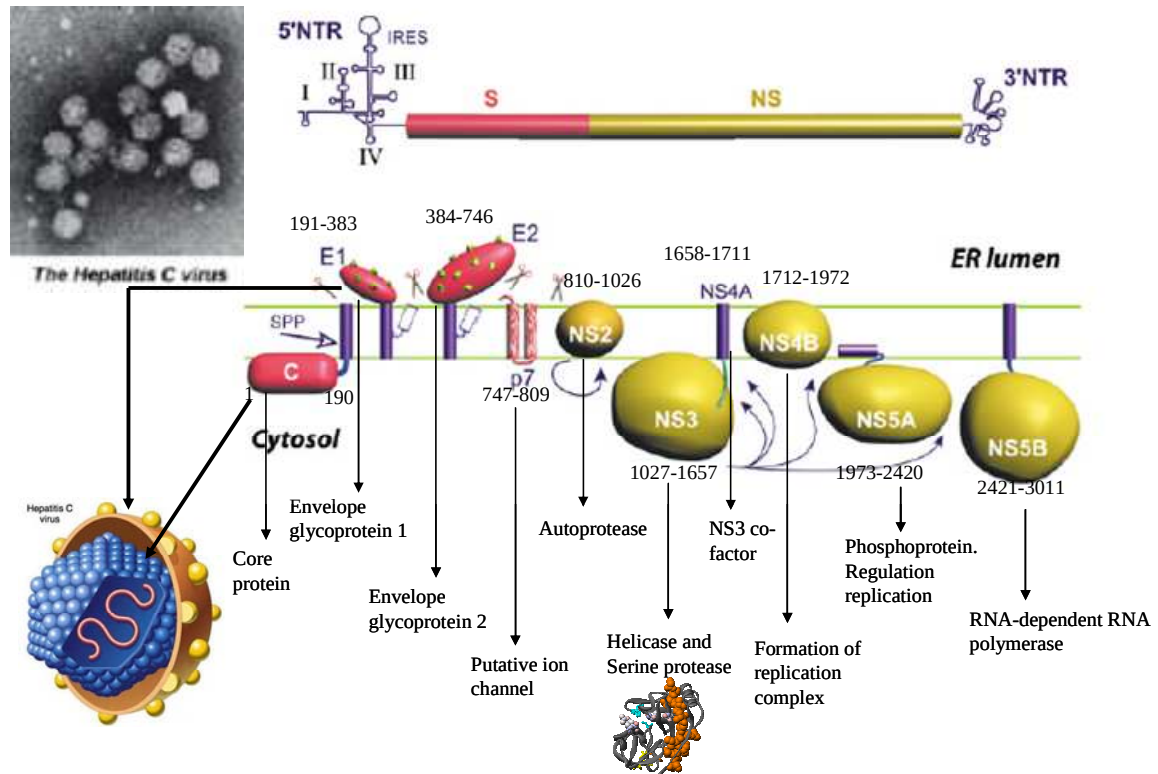


Figura 1. Genoma del VHC.

2.2 Variabilidad genética del VHC. Genotipos y subtipos

Los virus de ARN se caracterizan por tener una elevada variabilidad genética. El descubrimiento de un creciente número de variantes del VHC llevó a una normalización de la nomenclatura y estas variantes fueron clasificadas en genotipos y subtipos. En la actualidad las variantes del VHC se clasifican en 6 genotipos definidos mediante análisis filogenético. Los genotipos se enumeran del 1 al 6 y los subtipos se designan como a, b, c, y así sucesivamente (12). Los genotipos 1a, 1b, 2a y 2b, y 3a se distribuyen de forma global y representan la mayoría de las infecciones en todo el mundo. El G1b es el genotipo más común en el mundo y predomina en Norte de África, Sur y Este de Europa (13), México, Venezuela, Argentina, Japón, Corea del Sur, China y Taiwán. El genotipo 1a

predomina en América del Norte, Norte Europa, Perú y Puerto Rico. Los genotipos 2a y 2b son más comunes en América del Norte, Europa y Japón, y el G2c se encuentra en el Norte de Italia y en Córdoba (Argentina) (14-16). El G3 es altamente prevalente en el sudeste de Asia, Australia y Nueva Zelanda (17;18). El subtipo G3a, prevalente en el sudeste de Asia, ha ido en aumento en Europa y EE.UU, y se ha asociado a transmisión por drogadicción por vía intravenosa. Los genotipos 4, 5 y 6 están más restringidos geográficamente. El genotipo 4 prevalece en Oriente Medio y África, aunque también se ha detectado en aumento en Europa (Italia, España, Grecia, Suiza, Portugal, Francia, Suecia, Polonia, Alemania y Reino Unido), EE.UU. y el sur de la India, reflejando patrones de inmigración en estas áreas. El G5 se encuentra en Sudáfrica y el G6 restringido al sudeste de Asia, ha sido aislado en Vietnam, Hong Kong, Sur de China, Laos, Tailandia y Myanmar.

2.3 Dinámica de los virus de ARN. Quasiespecies

Los mecanismos de variación en los virus de ARN incluyen mutación, recombinación homóloga y no homóloga y reordenamiento génico en virus de genoma segmentado.

La causa principal de variación en los virus de ARN es la introducción de mutaciones, y ello es debido a:

1-Limitada fidelidad de copia (incorporación de nucleótido correcto) de las ARN polimerasas y retrotranscriptasas.

2- Ausencia de actividad correctora 3'-5' exonucleasa de la ARN polimerasa dependiente de ARN (19;20), durante la replicación. La tasa de mutación (sustitución/nucleótido/ciclo de replicación) espontánea es, pues, muy elevada. Los valores de tasa de mutación en la replicación de virus de ARN se han estimado del orden de 10^{-3} - 10^{-5} (20-22). En el caso del VHC la tasa de fijación de mutaciones (sust/nt/año) se ha calculado que es del orden de 1.5×10^{-3} (23). La tasa de fijación o acumulación de mutaciones, no obstante, no es la misma a lo largo del genoma. La región más heterogénea se encuentra entre las regiones estructurales E1 y E2,

donde se reconocen dos regiones hipervariables (HVR1 y HVR2). Otras regiones como core y NS3 están más conservadas y la región 5'UTR es la región con menor tasa de mutación, debido probablemente a las constricciones o limitaciones a aceptar mutaciones en una región con elevada estructura secundaria y presencia de señales de traducción y transcripción. Estudios cinéticos han determinado que se producen del orden de 10^{12} nuevos viriones de VHC por día en un individuo infectado, y se ha calculado que la vida media del VHC es de 2,7 horas (24). La consecuencia de esta cinética de replicación es que la población viral en un individuo infectado consiste en una mezcla de genomas diferentes entre sí, pero que están relacionados (98-99% homología), y que pueden cooperar, presentando una estructura que se denomina *quasiespecies*. La composición de las quasiespecies está sujeta a cambios continuos siguiendo los principios de evolución darwiniana: reproducción con variación genética, competición, selección y cooperación e interacción. El VHC presenta una estructura en quasiespecies (25;26).

Tabla 1. Tasas de mutación de diferentes virus de ARN

AUSENCIA DE MECANISMO DE CORRECCIÓN DE ERRORES

Tasa de acumulación (fijación) de mutaciones (tasa de evolución) = sustituciones/base/año

VHA: 1×10^{-3} - 10^{-4}

VHB: 1.4 - 7.9×10^{-5}

VHC: 1.5×10^{-3}

VHD: 3.0 - 5.0×10^{-3}

VHE: 1.40 - 1.72×10^{-3}

Tasa de mutación = sustitución/base/ciclo de replicación

VIH 3.4×10^{-5}

Los virus de ARN explotan diversos mecanismos de variación genética para asegurar su supervivencia (apartado 2.1). La dinámica viene marcada, en primer lugar, por la continua aparición y selección de mutantes que configuran la quasiespecies, evolucionando en un individuo infectado, y su transmisión de paciente a paciente usando diferentes rutas de infección. La segunda es la selección de diferentes cepas durante la evolución en poblaciones geográficas o epidemiológicas distintas (12). El modelo de quasiespecies fue originalmente

desarrollado como una teoría determinística de la evolución molecular propuesto por Manfred Eigen y Peter Schuster para explicar la organización y adaptabilidad de los replicones primitivos a las primeras etapas del desarrollo de la vida en la tierra (27). Desde entonces los virus han estado considerados por los biólogos teóricos como sistemas de modelos experimentales que estudian la dinámica de las quasiespecies, y los virólogos han encontrado soporte para interpretar observaciones experimentales de quasiespecies en infecciones en cultivos celulares y en sistemas *in vivo* (28). Esta distribución de mutantes se demostró experimentalmente primero con el bacteriófago Q β durante su replicación en el huésped de *E. coli* (21;29), y posteriormente se ha evidenciado para muchos virus de ARN, incluyendo patógenos de animales y humanos, el virus de la estomatitis vesicular y el “foot-and-mouth disease virus” (FMDV), el virus de la inmunodeficiencia humana tipo 1, y el virus de la hepatitis C (25;30-34).

Las quasiespecies virales se representan mediante una secuencia consenso y un espectro de mutantes (Figura 2). La secuencia consenso es la que presenta en cada posición el nucleótido más frecuente en el conjunto de genomas de la población y puede no estar representada en el espectro de mutantes de la población. Las quasiespecies estables suelen estar dominadas por una secuencia más frecuente en el espectro de mutantes, de mayor eficacia biológica, que se denomina secuencia máster y que puede o no coincidir con la secuencia consenso (35).

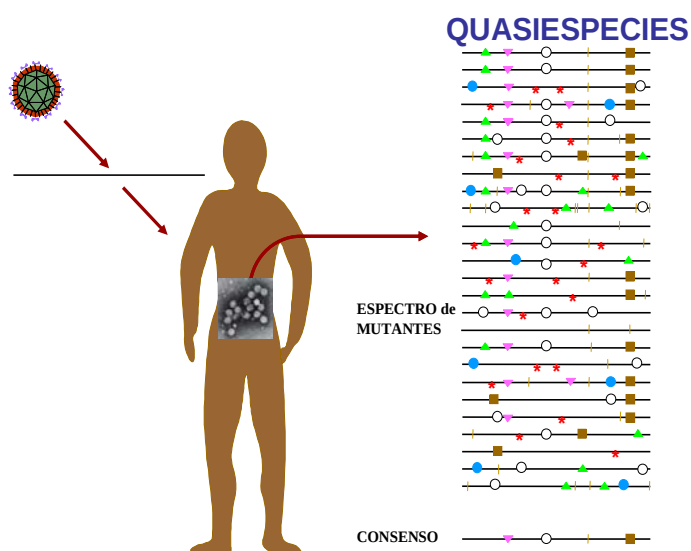


Figura 2. Esquema del concepto de quasiespecies y espectro de mutantes que la componen. Los genomas individuales se representan por rayas horizontales y los símbolos sobre las rayas representan mutaciones. Al conjunto de mutantes se le denomina espectro de mutantes de las quasiespecies.

Las quasiespecies están sujetas a continuos cambios, debido a procesos de selección competitiva de mutantes (26;36-38) y a la cooperación de diferentes variantes con diferentes fenotipos y capacidad de replicación (“fitness”) (39). La presencia simultánea de múltiples variantes genómicas (y la elevada tasa de generación de nuevas variantes) supone una ventaja adaptativa porque permite la rápida selección de mutantes con mejor “fitness”, para cambios en el medio de replicación, y la adaptación a infectar nuevos individuos y nuevos tipos celulares.

Los procesos de recombinación (especialmente en virus ARN de polaridad positiva) proveen a los virus ARN de un mecanismo adaptativo adicional a las altas tasas de mutación (40;41) y existen indicios de que también actúan en VHC (42;43).

La consiguiente variabilidad genética en las poblaciones virales conduce a un proceso evolutivo de selección natural darwiniana, en el que la diversificación de las poblaciones virales conlleva a la competición y a la cooperación (39). En este sentido se propone que la unidad de evolución sea la quasiespecies y no el individuo (44). Además de la continua aparición y selección de mutantes en un individuo, la transmisión de paciente a paciente usando diferentes rutas de infección es un mecanismo de variación genética y de adaptación (12).

Fitness es un término complejo que describe la adaptabilidad replicativa de un organismo a su medio. Para los virus de ARN, una aproximación experimental a “fitness” es la relativa habilidad para producir progenie infecciosa estable en un determinado medio ambiente (45-47). Una de las consecuencias de la estructura en quasiespecies de los virus de ARN es un efecto importante del tamaño de la población viral en la variación de “fitness” (48). Las poblaciones de virus de ARN sujetas a infecciones seriadas masivas en un cultivo celular bajo condiciones ambientales constantes tienden a ganar “fitness” en el ambiente, como resultado de la selección positiva de las variantes más eficaces que se generan durante la replicación, permitiendo la ganancia de la eficacia biológica (49). En contraste, sucesivos cuellos de botella poblacionales, en los que un sólo o varios genomas inician la infección de un modo repetitivo, ocasionarán descensos en la eficacia biológica, debido a la acumulación de mutaciones deletéreas (49-52). La

transmisión por cuello de botella alternando con transmisión poblacional in vivo puede haber condicionado la evolución del VHC in vivo.

2.4 Infección por VHC

La prevalencia del VHC en la población general es de alrededor del 2.3% (160 millones de personas infectadas). La OMS ha declarado el VHC un problema de salud global. Después de la infección aguda, alrededor del 50%-80% de los sujetos desarrollan hepatitis crónica con persistencia viral y con riesgo de desarrollar cirrosis hepática y/o carcinoma hepatocelular.

La curación espontánea del VHC ocurre aproximadamente en el 15% de los pacientes después de la infección primaria (53). Después de la infección por VHC, el curso natural de la infección por hepatitis C hacia curación espontánea o cronicidad depende de factores virales (huésped) y factores genómicos del paciente (Figura 3).

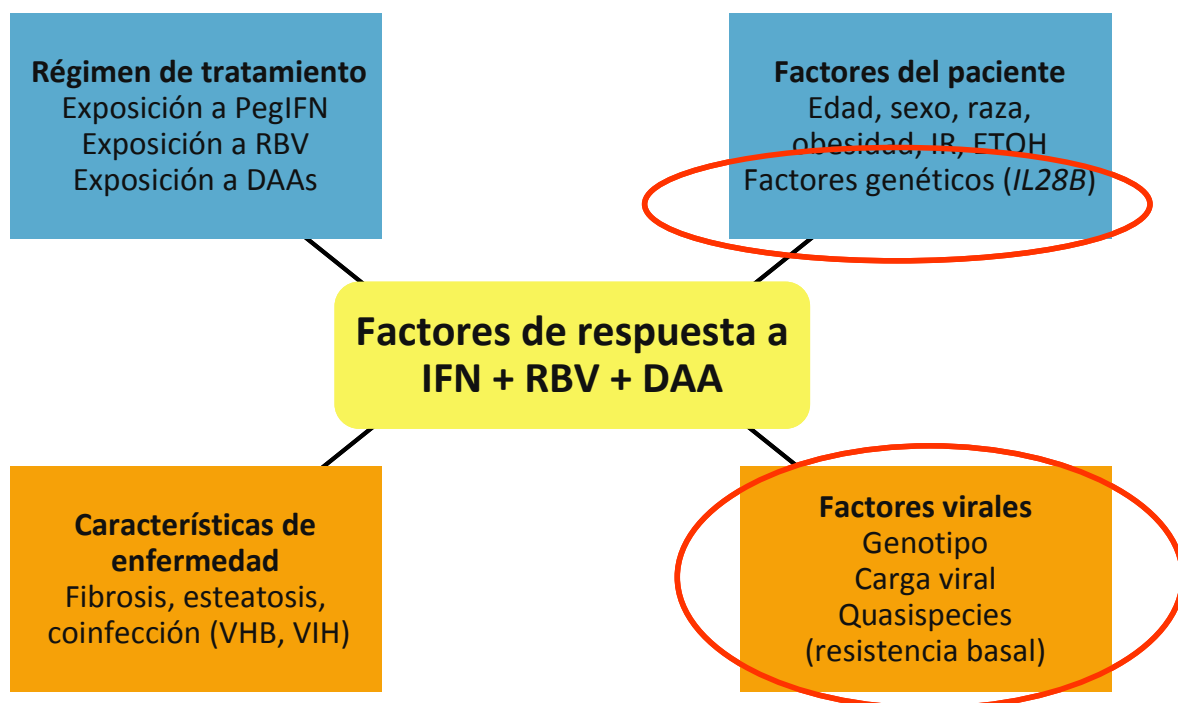


Figura 3. Factores relacionados con la evolución de la infección por VHC

2.5 Factores virales que influyen en la evolución de la infección por VHC hacia cronicidad o curación

Las tasas altas de mutación y la dinámica de quasiespecies afectan a la acumulación de mutaciones, las variaciones de la “fitness” y la supervivencia del virus. Se ha propuesto que existen 3 parámetros críticos involucrados en la persistencia del virus: la tasa de mutación, la tasa de replicación viral y la carga viral (número de partículas infecciosas en una población viral).

La teoría evolutiva predice que las tasas de mutación altas se ven favorecidas en un entorno dinámico, y las tasas de error virales pueden haber sido optimizadas por la selección natural. Respecto a los virus ARN, la fidelidad de replicación baja genera una población diversa de variantes. Si bien las variantes están en condiciones en general con una menor “fitness”, pueden dominar rápidamente si ocurre un cambio repentino en el ambiente, como por ejemplo, una respuesta inmunitaria contra la población mayoritaria. Por el contrario, una población homogénea, generada por una alta fidelidad de replicación, carece de flexibilidad y ésta podría ser menos exitosa en el entorno dinámico del ambiente del huésped (54).

La distribución en quasiespecies tiene numerosas implicaciones biológicas tales como el establecimiento de infección persistente por selección de mutantes de escape frente a los anticuerpos neutralizantes y linfocitos T, la selección de mutantes resistentes a fármacos antivirales, los cambios en la virulencia o patogenicidad del virus, y selección de mutantes de escape frente a vacunas, entre otras. La rápida evolución de los virus de ARN conlleva serios problemas para el desarrollo de vacunas.

2.5.1 Complejidad de la población viral

Se ha postulado que un aumento en la diversidad de la quasiespecies se asocia con una mayor virulencia (39). En este sentido, los pacientes infectados por el VHC que sufrieron poca evolución durante la fase aguda de la infección, se recuperaron completamente, mientras que los pacientes en los que la quasiespecies evolucionó rápidamente desarrollaron una infección crónica (55). En cuanto a la respuesta al tratamiento, se ha demostrado que cuanto más homogénea es la población mejor respuesta se consigue al tratamiento combinado pegIFN+RBV. Esto se demuestra porque el 80% de pacientes tratados en fase aguda (cuando se produce una homogeneización de la población) consiguen una RVS independientemente del genotipo, mientras que en fase crónica cuando la población viral es más compleja, la eficacia es menor y depende del genotipo, tan sólo se consiguen respuestas del 30-45% para genotipos 1 y 4, siendo del 80% para 2 y 3.

Además, el modo de transmisión también puede ser otro de los factores que determinen la transición de la infección aguda a persistente, debido a la transmisión más o menos limitada de un número físico de viriones.

2.5.2 Modo de transmisión del VHC

La transmisión del VHC a un nuevo paciente es una oportunidad única para identificar características virales asociadas con la evolución de la enfermedad. El modo de transmisión del VHC asociado a la cantidad de virus transmitidos podría ser uno de los factores clave para la transición de infección aguda a persistente. En primer lugar, cuando se produce la transmisión masiva de partículas virales, por ejemplo por transfusión o durante la recurrencia de la infección después de un trasplante hepático asociado a estado final de enfermedad hepática, la infección se convierte en persistente en el 80-100% de los casos respectivamente.

Por otro lado, en situaciones en que se transmite un inóculo pequeño (por exposición accidental a pinchazos con aguja infectada, relaciones sexuales, tatuaje, acupuntura o transmisión vertical madre-hijo) se han demostrado tasas medias de

transmisión del orden de 4-7% y la prevalencia de anticuerpos detectables anti-VCH es aproximadamente del rango entre 0,1% a 2,4% (56-60). Aún más interesante es que el 40% de parejas de portadores crónicos VHC+ tienen respuesta inmune CD4 de memoria específica contra NS3 en ausencia de evidencia clínica, molecular y serológica de exposición, lo cual sugiere que se ha producido una infección subclínica con eliminación completa del virus (61). Paralelamente, se ha detectado respuesta específica de células T contra proteínas no-estructurales del VHC en el 30% de individuos que han tenido contactos sexuales con pacientes con infección aguda C, y una vez más en ausencia de viremia detectable o seroconversión a anticuerpos. El riesgo menor de desarrollar persistencia viral después de la exposición a un inóculo viral pequeño, se ha relacionado con un fenómeno de cuello de botella genético (50;52;62-64). El cuello de botella genético definido como la transmisión de un número limitado de partículas virales, se relaciona con una reducción drástica y al azar de la diversidad viral. Estas partículas llevarán una cierta carga mutacional que puede limitar su capacidad de generación de variantes comparado con lo que ocurriría si se transmitiera una gran población (65). La carga mutacional y el escaso número de variantes, será un límite que tiene el virus para conseguir superar la respuesta inmune antiviral con la finalidad de persistir.

2.5.3 Evasión del sistema inmune

Algunos cambios fenotípicos de los virus están mediados por una sola sustitución de un nucleótido o aminoácido. Para tener valor adaptativo, un cambio genético mínimo no sólo debe de ser producido por la maquinaria de copia del genoma, sino que no debe suponer un coste de “fitness” de tal intensidad que no disminuya la proporción de mutantes por debajo de su nivel mínimo de umbral de expresión (66). Cuando se cumplen estos requisitos, el mutante estará representado en la quasiespecies con una frecuencia tal que las características fenotípicas asociadas con el mutante estarán expuestas a los eventos de selección.

Los mutantes más frecuentes de este tipo son: mutantes de escape a inhibidores o a respuesta específica de linfocitos T CD4+ y/o CD8+.

Los mutantes de escape a respuesta inmune del VHC, y su implicación en la persistencia viral, han sido descritos durante la infección de chimpancés. Estos estudios sugieren que una respuesta T CD8+ enfocada a pocos epítomos, junto con la aparición de mutantes de escape, promueve la persistencia del virus, mientras que una respuesta CD8+ amplia y temprana, favorece la eliminación del virus (67;68). La contribución de mutantes que eluden el reconocimiento de células T CD4+ “helper” o linfocitos T CD8+ citotóxicos (CTL) han sido descritos para numerosos sistemas de huésped-virus (69;70). La cinética de replicación viral en relación con el momento de la respuesta inmune del huésped puede ser crucial para la eliminación del virus. Esto significa que se produce una generación de variantes antigénicas del VHC que compiten entre ellas, y que la capacidad de estas variantes para convertirse en dominantes, depende de la generación e intensidad de las respuestas de las células T CD4+ y CD8+.

2.5.4 Resistencia a drogas antivirales

Las mutaciones de resistencia a drogas antivirales, supone un gran inconveniente en el tratamiento de la infección crónica por el VHC. La estructura en quasiespecies de los virus de ARN predice la presencia en cualquier paciente de todas las posibles variantes genómicas que portan sustituciones nucleotídicas simples en cada posición, condicionando la “fitness” relativa de cada mutante individual. La presencia en un huésped de todos los simples, dobles o triples mutantes depende de la carga viral (con 4×10^4 partículas todos los simples mutantes podrían estar representados, con 1.6×10^9 partículas todos los dobles mutantes y con 6×10^{13} partículas todos los simples, dobles y triples mutantes). Con niveles elevados de replicación del VHC en un paciente no tratado con antivirales (10^{12} partículas virales por día), mutaciones puntuales en cada posición del genoma surgen al menos una vez al día, de este modo, todos los mutantes resistentes son constantemente generados.

Los mutantes resistentes a antivirales son menos eficientes en la replicación que los virus “wild-type (wt)” y en la ausencia de terapia antiviral, representan mutantes minoritarios. Durante el tratamiento con antivirales, los mutantes son seleccionados y surgen como componentes dominantes de la población viral que conllevan al fallo del tratamiento. Además, las mutaciones compensatorias seleccionadas en el mismo genoma contribuyen a la recuperación de la “fitness” (71). Cuando se interrumpe el tratamiento, si dichas mutaciones son menos competitivas que otras, podrían ser seleccionadas negativamente y quedar relegadas a una población minoritaria, generando una memoria en la quasiespecies (72). En VHC se ha observado la persistencia de las mutaciones de resistencia a inhibidores de NS3 incluso aproximadamente a los 5 años (73) después de parar el tratamiento con desaparición del mutante. El re-tratamiento de estos pacientes con el mismo antiviral podría suponer la reaparición del mutante con mayor celeridad.

2.6 Factores del paciente que influyen en la evolución de la infección por VHC hacia cronicidad o curación

La evolución de la infección por VHC varía sustancialmente entre los diferentes individuos. Los factores genéticos del hospedador también tienen un papel determinante en dicha evolución, y son: factores personales (edad, peso, sexo, etnia) y de enfermedad (fibrosis, obesidad, esteatosis hepática, LDL colesterol o resistencia a insulina) junto con la respuesta inmune e inflamatoria dirigida por polimorfismos de HLA, heterogeneidad celular, variaciones en la temperatura, etc., que pueden influir en el resultado de la infección hacia cronicidad o resolución espontánea (74-77). Recientemente, se han descrito polimorfismos en el ADN genómico directamente asociados a RVS y a la resolución espontánea de la infección aguda por el VHC (78).

2.6.1 [*IL-28B*](#)

En el año 2009 se realizó un estudio de asociación de genoma completo (genome wide association study-GWAS-) en un grupo de más de 1600 pacientes VHC+ de diferentes etnias, que recibieron tratamiento con interferón pegilado (IFN)- α 2a o 2b y ribavirina (RBV), para identificar variantes genéticas predictoras de respuesta al tratamiento en la hepatitis C crónica (79). Se identificó un polimorfismo de nucleótido único (SNP) en el cromosoma 19 (19q13.2), rs12979860, fuertemente asociado con respuesta virológica sostenida (RVS). Este SNP está localizado 3kb “upstream” del gen IL28B, que codifica el interferón λ de tipo III (IFN- λ 3). La familia de los interferones λ IFN- λ 1, 2 y 3 (también denominados IL-29, IL28A e IL-28B) son un grupo de citoquinas antivirales relacionados con la familia de IFNs tipo I e IL-10. El genotipo C/C en rs12979860 fue asociado a RVS con una tasa 2.5 veces superior (dependiendo de la etnia) que el genotipo T/T. El genotipo C/C favorable fue más común en Caucásicos que Afro-Americanos. También se describió en un estudio que la misma variante IL-28B también estaba asociada con una mayor probabilidad de curación espontánea tras la infección aguda por virus (80). Otro estudio GWAS desarrollado en pacientes japoneses con infección por VHC de genotipo 1 puso de manifiesto otro SNP (rs8099917-localizado 8kb “upstream” del gen IL28B) como la variante más fuertemente asociada a RVS (81). Este mismo SNP también fue identificado en una cohorte de 293 pacientes australianos infectados con VHC de genotipo 1 (82).

2.6.2 [*Sistema inmune*](#)

El perfil de respuesta inmune frente a VHC ha sido ampliamente estudiado en humanos y chimpancés, y ha permitido determinar la correlación entre el tipo y vigor de la misma y la persistencia o resolución de la infección. En la infección aguda, la replicación viral es detectable durante la primera semana, alcanzando niveles máximos 4 semanas después de la exposición. La respuesta de las células CTL se desarrolla por completo coincidiendo con los valores máximos de transaminasas hepáticas. Estas células efectoras CTL tempranas no producen IFN- γ ,

pero son eficientes en la destrucción de hepatocitos infectados. En muchos casos una respuesta Th1 CD4 multiespecífica aparece posteriormente y coincide con una fuerte disminución del nivel de ARN del VHC, normalización de ALT, y un cambio de CTLs hacia una población productora de IFN- γ . Los perfiles de activación hepática en fase aguda en chimpancé han demostrado una correlación directa entre la “upregulation” de genes inductores de IFN- γ y la inhibición de la replicación viral (83;84). Entre los individuos que resuelven la infección, las respuestas específicas se mantienen indefinidamente, mientras que los anticuerpos declinan progresivamente, y eventualmente, desaparecen décadas después de la exposición (85). Esto explica la detección frecuente de la respuesta inmune celular en individuos VHC seronegativos (86), y especialmente entre los cónyuges de los portadores crónicos (61). Otro factor del huésped que influye en el resultado de la infección es la respuesta inmune de las células T ejercida durante la infección aguda. Una respuesta inmune sostenida y multiespecífica de células T CD4+ y CD8+ tipo Th1 (producción de IFN- γ e IL-2) dirigida contra las proteínas no estructurales del VHC, principalmente contra la proteína NS3, está fuertemente asociada con la eliminación del virus con éxito durante la fase aguda de la hepatitis. En un estudio se identificaron 13 epítomos de células CD4 dentro de la región NS3-NS4 que fueron reconocidos por >30% de pacientes con infección aguda o resuelta por VHC (87). Si esta respuesta es limitada en el tiempo con cambio hacia respuesta tipo Th2 caracterizada por la producción de IL-10, IL-4 y respuesta de anticuerpos, va seguida por el rebrote viral y persistencia (88-94).

2.6.3 Otros factores

Una serie de estudios han demostrado que polimorfismos en el complejo mayor de histocompatibilidad (CMH) de clase II están relacionados con la evolución de la infección por VHC. Los genes del CMH de clase II controlan la respuesta inmune a través del reconocimiento de péptidos antigénicos. Las moléculas HLA son glico-proteínas de la superficie celular que se unen a péptidos y lo presentan a células CD4 T-“helper” durante la respuesta inmune. En humanos

algunos alelos de MHC clase II se han asociado con persistencia o resolución de la infección por VHC (95), en concreto, los haplotipos HLA clase II DRB1*0101, DRB1*1101, DQB1*0301, DQB1*0302 y DQB1*0501 se han relacionado con la resolución espontánea de la infección por VHC (96-101). Con respecto al haplotipo HLA-DRB1*1501, se ha identificado en la literatura un epítipo Th1 (1388-1407) restringido a este HLA y mutantes de escape en esta región, pueden atenuar o inhibir la proliferación de las células T (89;102). Otros haplotipos de HLA de clase II como DQB1*02 y DQB1*06 han sido relacionados con persistencia de la infección (101).

En los últimos años, factores del huésped han sido fuertemente asociados con el resultado del tratamiento estándar del VHC (SOC). Por lo tanto, para predecir con exactitud la respuesta al tratamiento estándar se requiere la identificación de estos factores asociados con la respuesta al tratamiento, para diseñar estrategias individualizadas de tratamiento. La disponibilidad en la actualidad de plataformas de alto rendimiento de secuenciación de poblaciones virales y de polimorfismos del huésped podría facilitar la optimización del tratamiento.

2.7 Plataformas de secuenciación “high-throughput”

En 1990 se firmó el “Proyecto Genoma Humano”. Más tarde, en 1995, se publicó el primer genoma secuenciado de un organismo de vida libre, *Haemophilus influenzae*. En el año 2001, años antes de lo que se tenía previsto, se publicó el primer borrador del genoma humano, que costó casi 3000 millones de dólares para leer 3000 millones de nucleótidos (1\$/nt). La necesidad de reducir costes de tiempo, trabajo y económicos, llevó al diseño de los secuenciadores de segunda generación y tercera generación, capaces de generar cientos de miles de reacciones en paralelo (secuenciadores de verdadero alto rendimiento [*high-throughput*]) gracias a la inmovilización de las reacciones en una superficie sólida.

Actualmente se encuentran en el mercado cinco tipos de secuenciadores de ADN de alto rendimiento (Tabla 2).

Tabla 2. Comparación de las distintas plataformas de secuenciadores

Equipo	Compañía	Método de secuenciación	Máximo Longitud de lectura (pb)	Número de reads (x10 ⁶)
GS-FLX (454)	454 Life Sciences/Roche Diagnostics	emPCR/pirosecuenciación	1000	1
GS-Junior			400	0,1
Miseq	Illumina	captura	150	3,4
Genome Analyzer IIx			150	320
HiSeq 1000		sólida/terminadores reversibles	100	1500
HiSeq 2000			100	300
5500	Applied Biosystems	emPCR/pirosecuenciación	75	800
5500xl			75	1600
Ion Torrent		emPCR/secuenciación real time	>400	6,2
Pac Bio RS	Pacific Biosciences/Gen-Probe	molécula	1200	1200
		única/secuenciación real time		
Nanopore	Oxford Nanopore	molécula única		

La primera aproximación a la secuenciación masiva en paralelo está basada en la pirosecuenciación del ADN. Se trata de una técnica no fluorescente que mide la liberación de pirofosfato en una reacción de polimerización mediante una serie de reacciones enzimáticas acopladas que liberan un fotón de luz cada vez que se incorpora un nucleótido; producen una imagen que se analiza para proporcionar flujogramas que, una vez interpretados por el ordenador, devuelven las secuencias de nucleótidos. El primer modelo comercial, el GS20, era capaz de secuenciar hasta 20 millones de bases en 4 horas (1 millón de secuencias de 250 bases). Mejorando simplemente la cinética de la reacción química de secuenciación, en el año 2007 apareció el modelo GS-FLX, capaz de producir hasta 100 millones de bases en un tiempo similar y en 2009 el GS-FLX-Titanium que genera secuencias clonales de 400-600 bases, duplicando la efectividad del modelo anterior. En 2011 apareció el GS-FLX Titanium plus que proporciona secuencias clonales de un tamaño de 800-1000 bases. La longitud de las lecturas generadas se acerca a las obtenidas por el método de Sanger, lo que las convierte en ideales para la secuencia de genomas de *novo*. Al mismo tiempo, otras 2 compañías desarrollaron otro tipo de tecnologías para la secuenciación masiva en paralelo del ADN. El sistema Illumina (Illumina, Inc, San Diego, CA) utiliza un método basado en la polimerización del ADN, donde en primer lugar, se realiza un fraccionamiento del genoma, seguido de una

amplificación clonal de fragmentos de ADN sobre la superficie sólida de una celda de flujo. El ADN es secuenciado a través de la polimerización con deoxinucleótidos fluorescentes (103). La 3ª tecnología es SOLiD (Sequencing by Oligonucleotide Ligation and Detection) comercializada por Applied Biosystems (Applied Biosystems, Carlsbad, CA). Esta tecnología secuencia por ligación de octámeros marcados de secuencia conocida a la cadena de ADN, con la posterior detección de la señal fluorescente emitida tras cada ligación. El sistema SOLiD usa una PCR en emulsión para enriquecer el ADN clonal fijado a microesferas. Las esferas se transfieren a un chip de flujo, y el ADN se secuencia por ligación con cebadores fluorescentes. Estas 2 tecnologías generan millones de secuencias cortas (50-100 bases). Estos métodos son adecuados para resecuenciación cuando una secuencia de referencia está disponible (genoma humano o transcriptoma) y la longitud de lectura es suficiente para identificar un gen conocido. También se pueden utilizar para identificar sustituciones individuales en genes virales, tales como las asociadas con resistencia a drogas. Sin embargo, las lecturas cortas no se pueden utilizar para establecer “linkage” entre sustituciones ubicadas en la misma secuencia, que son requeridas para estudios de resistencia viral. Otras tecnologías en desarrollo incluyen al Ion Torrent (Applied Biosystems), en la cual la emulsión de PCR de los amplicones se carga en un chip semiconductor, la adición de nucleótidos durante la reacción de polimerización genera iones H^+ , que inducen cambios locales de pH que se detectan y se miden (104). La tecnología PacBioRS (Pacific Biosciences, GenProbe, Menlo Park, CA) en la cual el ADN sin amplificar de la librería de simple se transfiere a un sistema óptico que amplifica las señales fluorescentes formadas por cada moléculas de ADN clonal, procesado por miles de moléculas de polimerasa (105). El proceso más novedoso, y que está en desarrollo, tiene que ver con el uso de nuevos materiales (el grafeno) que utiliza la tecnología Nanoporo (Oxford Nanopore), cuyo objetivo es secuenciar moléculas individuales sin amplificar ni manipular. Esta técnica utiliza cargas eléctricas para impulsar en un chip de Silicon, cadenas de ácidos nucleicos a una velocidad constante a través de nanoporos de 50 nanómetros. Luego 2 electrodos miden la magnitud de corriente proveniente de cada base que pasa a través del nanoporo de manera que cada

nucleótido que atraviesa el nanoporo genera un cambio de corriente diferente. Estas tecnologías están en constante evolución y han revolucionado el concepto de secuenciación (106).

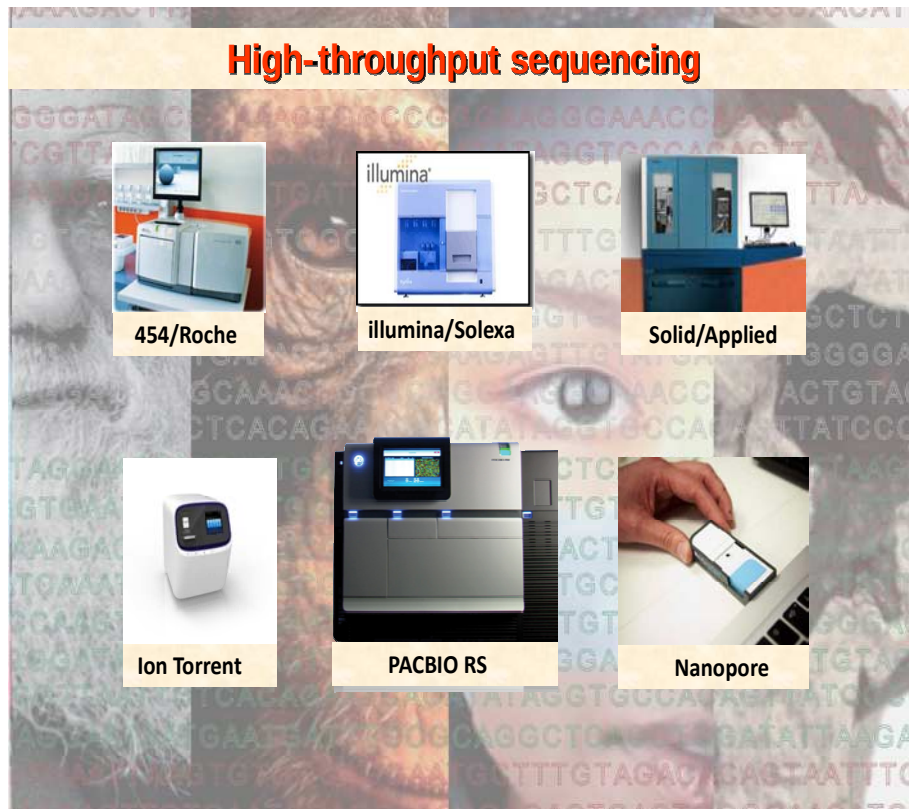


Figura 4. Diferentes técnicas de secuenciación de última generación

ESTUDIO 1

3 ESTUDIO 1: "EFFECT OF BOTTLENECKING ON EVOLUTION OF THE NONSTRUCTURAL PROTEIN 3 GENE OF HEPATITIS C VIRUS DURING SEXUALLY TRANSMITTED ACUTE RESOLVING INFECTION"

3.1 Hipótesis:

Las parejas sexuales de pacientes infectados con el VHC, presentan a menudo respuestas específicas de células T en ausencia de seroconversión, sugiriendo la resolución espontánea de la infección. El riesgo bajo de persistencia del virus después de la exposición sexual podría estar relacionado con la transmisión de un número limitado de partículas virales (*"genetic bottleneck"*).

3.2 Objetivo:

El objetivo fue el estudio de la diversificación de la secuencia del gen NS3 en un paciente con infección crónica y las características de las quasiespecies virales antes y después de la transmisión en un hospedador nuevo, su pareja, que desarrolló una infección aguda con resolución espontánea, mediante secuenciación clásica por Sanger.

3.3 Resumen:

Alrededor del 20% de los pacientes expuestos al VHC resuelven espontáneamente la infección. La recuperación es más común entre los pacientes con hepatitis aguda y que se han infectado por inóculos pequeños, como los producidos por pinchazos accidentales, uso de drogas intravenosas y por transmisión sexual (56-58;60). Así, una complejidad menor de la quasiespecie viral resultante de inóculo pequeño que ha pasado por un fenómeno de "cuello de botella genético" se ha asociado a una mayor tasa de resolución espontánea. La transmisión sexual del VHC es un tema que genera controversia. Sin embargo, evidencias de estudios de parejas regulares de pacientes por VHC han demostrado que existe un riesgo bajo de

transmisión por VHC, y una posible razón podría ser la presencia de ARN del VHC en el semen en un nivel reducido ($<200\text{copias/ml}$ -**(107)**). En el estudio 1 que se adjunta, se analizó un caso en que el VHC se transmitió de un paciente con infección crónica por VHC (denominado paciente A) a su pareja (paciente B), después de la eliminación de un dispositivo intrauterino, causante de lesiones vaginales y que probablemente facilitó la transmisión del VHC del semen a la sangre. A partir de las muestras de suero de ambos pacientes, obtenidas lo más cerca posible del momento de la transmisión (septiembre/octubre 1998), se analizaron 12 clones tanto del paciente A (VHC crónico) como del paciente B (infección aguda) con la técnica de clonación-secuenciación por Sanger. La comparación de los parámetros de quasiespecies en el aislado del paciente A comparado con el aislado de su pareja (paciente B), reveló una heterogeneidad elevada y una proporción baja de mutaciones no-sinónimas en la muestra origen. El árbol filogenético consenso construido con secuencias nucleotídicas promedio consecutivas del paciente A y 24 clones de ambos pacientes mostró que todas las secuencias del paciente B se agruparon con el clon A3 (procedente del paciente A-VHC crónico), esto sugiere un fenómeno de cuello de botella durante la transmisión. La drástica reducción en el número y diversidad de las partículas infecciosas que originan la infección después de la exposición sexual al VHC, junto con la reducida adaptabilidad de los genomas de muestras aleatorias con mutaciones deletéreas, podrían ser causantes de la limitación de la generación de mutantes capaces de superar la presión selectiva ejercida por la respuesta inmune innata y adaptativa en el nuevo hospedador. La caracterización de las poblaciones virales en el origen de la infección, y también durante la progresión hacia cronicidad es una información fundamental para la comprensión de la patogénesis y el diseño de vacunas. En este trabajo se midió la tasa global de fijación de mutaciones o tasa de evolución a partir de la secuencia directa del producto de PCR de 5 muestras secuenciales de suero del paciente A (1993, 96, 97, 98, 2001). La tasa global de fijación de mutaciones o tasa de mutación, estimada por el análisis de regresión de las secuencias consenso de más de 8 años, fue de $1,5 \times 10^{-3}$ mutaciones/nt/año, una tasa relativamente estable con pequeñas fluctuaciones entre las muestras. Las limitaciones técnicas en el estudio de poblaciones complejas en el tiempo en que se

inició el trabajo y la reciente aparición de las técnicas de “high-throughput” han inducido a estudiar con más profundidad este trabajo que se presenta en el estudio 3.

Effect of Bottlenecking on Evolution of the Nonstructural Protein 3 Gene of Hepatitis C Virus during Sexually Transmitted Acute Resolving Infection

Josep Quer,^{1*} Juan Ignacio Esteban,¹ Joan Cos,² Sílvia Sauleda,² Laura Ocaña,² María Martell,¹ Teresa Otero,¹ Maria Cubero,¹ Eduard Palou,² Pedro Murillo,^{1†} Rafael Esteban,¹ and Jaume Guàrdia¹

Liver Unit, Department of Internal Medicine, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain,¹ and Banc de Sang i Teixits, Servei Català de la Salut, Barcelona, Spain²

Received 21 July 2005/Accepted 20 September 2005

Sexual partners of patients infected with the hepatitis C virus (HCV) often have detectable HCV-specific T-cell responses in the absence of seroconversion, suggesting unapparent, spontaneously resolving infection. To determine whether differences in the evolutionary potential of bottlenecked inoculum may explain the low rate of HCV persistence after sexual exposure, we have investigated changes in the entire HCV nonstructural 3 (NS3) gene over time in a chronic carrier and compared his viral quasispecies with that of the acute-phase isolate of his sexual partner, who developed acute resolving hepatitis C. The overall rate of accumulation of mutations, estimated by regression analysis of six consecutive consensus NS3 sequences over 8 years, was 1.5×10^{-3} mutations per site per year, with small intersample fluctuations related to changes in environmental conditions. Comparison of quasispecies parameters in one isolate of the chronic carrier with those of the acute-phase isolate of the infected partner revealed a higher heterogeneity and lower proportion of nonsynonymous mutations in the former. All NS3 sequences from the acute-phase isolate clustered with a single sequence from the chronic isolate, despite complete HLA mismatch between the patients, suggesting bottlenecking during transmission. The low risk of viral persistence after sexual exposure to HCV may be related to the selection of a limited number of viral particles carrying a particular combination of mutations which may further limit the potential of a relatively homogeneous quasispecies to rapidly diversify and overcome the immune response of the exposed host.

About 20% of patients exposed to hepatitis C virus (HCV) spontaneously resolve their infection. Recovery appears to be more common among patients with clinically evident acute hepatitis (61) and in those exposed to small inocula such as those produced by accidental needle sticks, intravenous drug use, and sexual transmission (26, 44, 45, 52, 66). In fact, a significant proportion of individuals with these kinds of exposure develop HCV-specific cellular immune responses in the absence of detectable viremia or antibody seroconversion (3, 29, 31, 57), and experimental evidence that low infective doses induce T-cell responses has been reported for the chimpanzee (58). Although the precise determinants of viral clearance or persistence are not well understood, a significant association has been found between a broad and sustained HCV-specific T-cell response and viral clearance in acute hepatitis C. At the CD4 T-cell level, such a response is directed especially against nonstructural proteins, and highly conserved immunodominant epitopes within nonstructural protein 3 (NS3) have been

identified, suggesting a key role of NS3-specific CD4⁺ responses and spontaneous recovery (5, 10, 25, 51).

HCV is an enveloped, single-stranded, positive-sense RNA virus belonging to the *Flaviviridae* family, genus *Hepacivirus*. Its 9.6-kb genome encodes a single polyprotein of about 3,011 amino acids, which is cleaved into three structural and six nonstructural proteins. The nonstructural proteins NS2 to NS5B are involved in polyprotein processing and viral replication. The product of the NS3 gene is a protein of 67 kDa which includes two domains, a serine proteinase that comprises the 189 N-terminal amino acids and a helicase-nucleoside triphosphatase domain spanning the 442 C-terminal amino acids (43). Like most RNA viruses, HCV evolves rapidly due to high mutation rates and high-level viral replication through an error-prone RNA polymerase without proofreading capacity. Consequently, in the infected individual, the viral population is composed of a complex mixture of different but closely related genomes known as quasispecies (37), whose shape is subject to continuous changes due to competitive selection of continuously arising mutants (13, 14, 19, 20, 27, 28, 59). The quasispecies nature of HCV is thought to play a central role in the establishment of persistence overwhelming innate and adaptive immune responses during acute infection. Once persistence is established, spontaneous recovery does not occur and eradication requires effective antiviral therapy. Among the nu-

* Corresponding author. Mailing address: Liver Unit, Department of Internal Medicine, Hospital Universitari Vall d'Hebron, Pg Vall d'Hebron 119-129, 08035 Barcelona, Spain. Phone: 34 934894028/4034. Fax: 34 934894032. E-mail: jquer@ir.vhebron.net.

† Present address: Departament de Biologia Molecular, Laboratori Cerba Internacional, Sabadell, Spain.

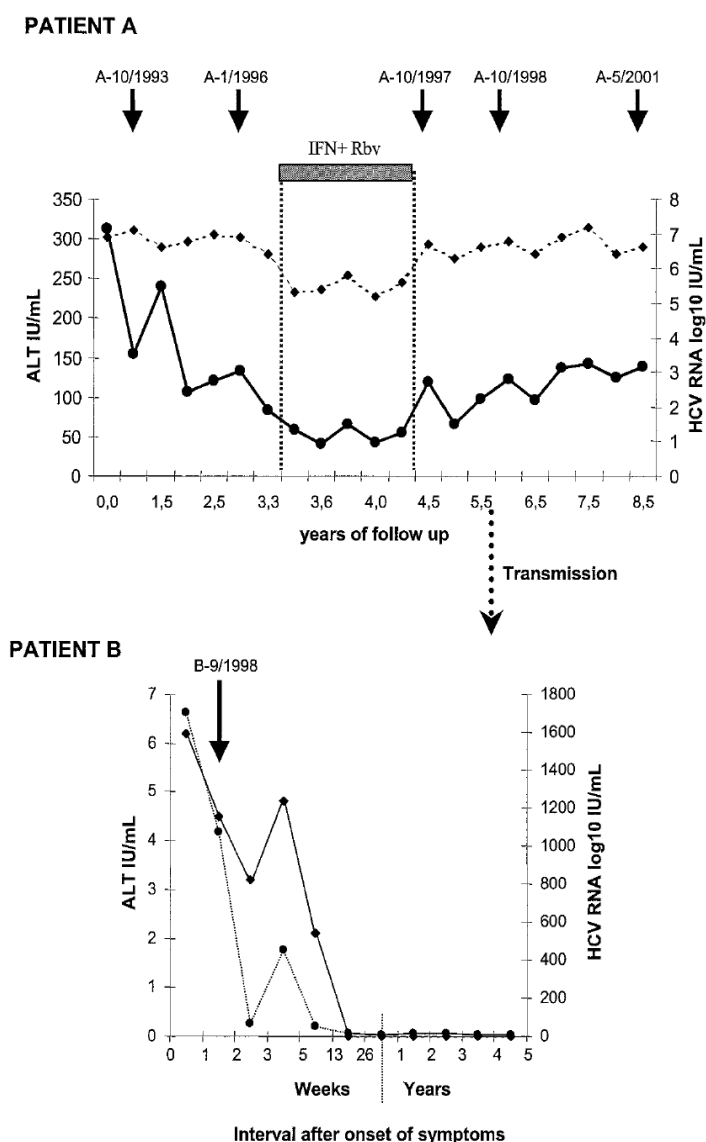


FIG. 1. Representative changes in ALT (solid line) and HCV viral load (dashed line) in patients A and B. Arrows indicate samples obtained for sequencing. The gray area indicates the antiviral treatment period for patient A (February 1996 to March 1997). IFN+Rbv, interferon plus ribavirin.

merous implications of the quasispecies structure for the biology of RNA viruses, one predicts relevant effects of genetic bottlenecks. Because of the extreme heterogeneity of the viral population, especially during persistent infection, any genome sampled at random from a mutant spectrum is likely to harbor deleterious mutations relative to the consensus, thus generating progeny with decreased fitness (12, 17, 21). The observation of fitness decrease after bottlenecks agrees with a genetic principle known as “Muller’s ratchet” (38). Muller’s ratchet describes the accumulation of deleterious mutations and overall fitness loss that occurs after repeated genetic bottleneck, and its

operation has been demonstrated in several viral species (6, 7, 12, 18, 22). It has been recently reported that in addition to causing fitness loss, genetic bottlenecks may also compromise the capacity of a viral population to adapt (39). In this regard, the overall decrease in mutant frequency associated with bottlenecks might limit the potential of the randomly sampled virus to overcome innate and adaptive immune responses in the new host and its ability to establish persistence.

In the current study, we have examined the sequence diversification of the NS3 gene in a chronically infected patient and the characteristics of the viral quasispecies before and after

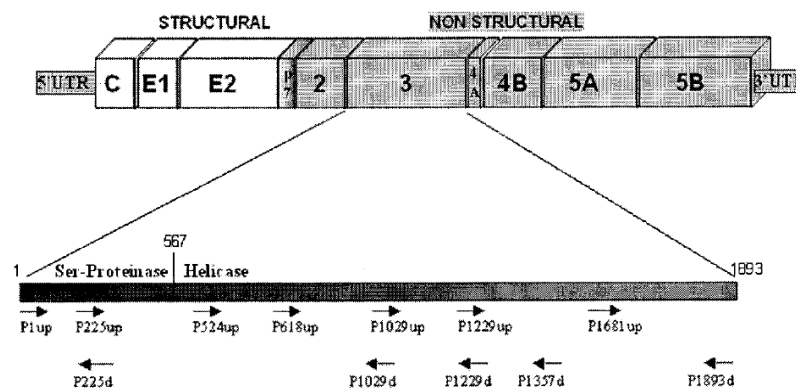


FIG. 2. Fragment of the HCV genome investigated in the study and locations of the PCR primers (arrows) used for direct sequencing of PCR products or clones. UTR, untranslated region.

transmission into a new host who developed acute resolving infection.

MATERIALS AND METHODS

Patient A was a 28-year-old man with a history of intravenous drug use who was found to be anti-HCV positive at the time of a volunteer blood donation in 1993 and has been prospectively followed since that time. He had persistent HCV infection with a genotype 1a and a viral load of 10^6 IU/ml. He had persistently elevated alanine aminotransferase (ALT) levels (mean, 178 IU/ml), and liver biopsy showed moderate chronic active hepatitis. He received treatment with interferon and Ribavirin for 12 months (between February 1996 and March 1997), with a transient biochemical but no virological response. A repeat liver biopsy performed in March 1998 showed no change in inflammatory or fibrosis scores. As shown in Fig. 1, five sequential serum samples obtained between October 1993 (A-10/1993) and May 2001 (A-5/2001) were used for direct PCR sequencing, and one sample (A-10/1998) was used for sequencing multiple clones. The HLA type II haplotype of patient A was DR1, DR15 (DRB1* 0102, 1501); DR51 (DRB5*), DQ5, DQ6 and type I A2, A28 (A*0201, 6802); B7, B14 (B*0702, 1402); and Cw7, Cw8 (Cw*0702,0802).

Patient B, a 26-year-old woman, was the sexual partner of patient A, and had been a regular blood donor whose last donation in December 1997 was anti-HCV negative. In August 1998 she developed acute symptomatic hepatitis C (peak bilirubin, 7 mg/dl; peak ALT, 1,705 IU/liter) as evidenced by anti-HCV seroconversion and presence of high-titer HCV RNA subtype 1a. The episode occurred around 7 weeks after unprotected sexual intercourse with patient A, shortly after removal of an intrauterine device due to vaginal bleeding. We have previously reported molecular phylogenetic data demonstrating that patient A was the source of the infection (48). Patient B normalized ALT levels within 6 weeks and cleared HCV RNA 8 weeks after onset of symptoms. HCV RNA has remained undetectable and ALT levels strictly normal to date. As shown in Fig. 1, a sample obtained during the acute phase of symptomatic hepatitis in patient B (B-9/1998) was used for cloning and sequencing. The HLA type II haplotype of patient B was DR4, DR11 (DRB1* 0402, 1101); DR52 (DRB3* 02); DR53 (DRB4*); DQ3, DQ3; and type I A66, A69 (A*6601, 6901); B41, B55 (B*4102, 5501); and Cw1 (Cw*0102,1701).

RNA extraction and nested RT-PCR. RNA was extracted from 200 μ l of serum by using Amplicor HCV specimen preparation kit v2.0 HCV-PREP (Roche, Basel, Switzerland). The measures to prevent contamination suggested by Kwok and Higuchi were strictly applied (32).

For a long reverse transcription-PCR (RT-PCR) product, 1/20 of the extracted RNA (10 μ l) was first denatured for 5 min at 70°C and kept on ice until used for cDNA production. RT was carried out for 60 min at 42°C, followed by a 15-min incubation step at 70°C (GeneAmp 2700 PCR system; Applied Biosystems, Foster City, CA), using 50 U Moloney murine leukemia virus RT H- point mutation (200 U/ μ l; Promega, Madison, WI), 20 U RNase inhibitor (40 U/ μ l; Promega, Madison, WI), 10 mM each deoxynucleoside triphosphate (Roche, Basel, Switzerland), 20 pmol of antisense PCR primer MNS3dout (5'TAGGAT CCAGGCAGCGTTTGACAAGCCCGCA3'), and 1 $\times$ buffer from high-fidelity *Pfu* Turbo DNA polymerase (Stratagene, San Diego, CA) in a final volume of 20

μ l. Then, 80 μ l of PCR mix containing 1 $\times$ *Pfu* Turbo buffer, 20 pmol of sense primer MNS3upout (5'ATGAATTCAGATACCGCCGCGTGCGGTGACAT3'), and 2.5 U *Pfu* Turbo DNA polymerase were added to each tube. After an initial step of denaturation for 2 min at 95°C, five initial cycles of 30 s at 94°C, 30 s at 55°C, and 2 min at 72°C were done, followed by 35 cycles of 30 s at 94°C, 30 s at 60°C, and 2 min at 72°C and finishing with a single final step of 10 min at 72°C. A product of 2,399 nucleotides was obtained.

Five microliters of the obtained product was used for nested PCR using internal primers P1up (up) (5'ATGGATCCATGGCGCCCATACGGCGTACGCCCAG3') and P1893d (down) (5'GTAAGCTTTTACGTGACGACC TCCAGGTCAGCCGACAT3'). The nested PCR mix consisted of 1 $\times$ *Pfu* Turbo buffer, 10 mM each deoxynucleoside triphosphate, 20 pmol internal primers, and 2.5 U *Pfu* Turbo DNA polymerase in a final volume of 100 μ l. After a single denaturation step of 2 min at 95°C, we carried out 30 cycles of 30 s at 95°C, 30 s at 60°C, and 2 min at 72°C and then a final single step of 10 min at 72°C. The amplified product of 1,915 nucleotides, including the NS3 region of 1,893 nucleotides (631 amino acids) and two short sequences for restriction enzymes at the 5' end (BamHI) and the 3' end (HindIII), was analyzed on a 2% agarose gel stained with ethidium bromide. PCR products were purified using the GeneClean Turbo for PCR system (Qbiogene Montreal, Canada) for cloning and/or sequencing.

Cloning and sequencing. Nested PCR products from samples A-10/1998 and B-9/1998 were cloned into the pCR-4Blunt TOPO plasmid by using the Zero Blunt TOPO PCR cloning kit in TOP10 *Escherichia coli* cells (Invitrogen, Carlsbad, CA), taking advantage of the fact that the *Pfu* DNA polymerase is devoid of terminal deoxynucleotidyl transferase activity and generates blunt-ended PCR products exclusively.

Overlapping primers in both directions (up and down) (Fig. 2) were used to obtain the complete NS3 sequence of 1,893 nucleotides from either PCR products or cloned fragments. Consensus sequence was obtained by direct sequencing of the nested PCR products. We used the Abi Prism BigDye terminator version 1.1 kit (Applied Biosystems) and analyzed the product with the Abi Prism 310 genetic analyzer (Applied Biosystems, Foster City, CA).

Sequence nomenclature. Clonal sequences are identified as A or B plus clone number, with A corresponding to the chronic carrier and B to the acute resolving infected partner. Consensus sequences from direct sequencing of PCR products are specified as A or B plus the month and year when the corresponding sample was obtained.

Sequence analysis. Nucleotide sequences were aligned by means of the CLUSTALW program (24). Phylogenetic analysis was carried out using the PHYLIP package (23) as follows: a bootstrap analysis was performed with 10,000 replicates, and data were analyzed by parsimony and distance methods that generated 10,000 trees. The consensus phylogenetic tree was drawn using Tree-view v1.6.5 (50).

Evolution of NS3 consensus sequence. The rate of accumulation of mutations was calculated as the number of nucleotide or amino acid mutations fixed per total number of nucleotides or amino acids analyzed per unit of time (in years) by comparing each sequence with both the most ancient one (1993) and its immediate predecessor. In both cases, the rate of fixation of mutations was inferred after regression analysis. The slope reflects the unit increase in substi-

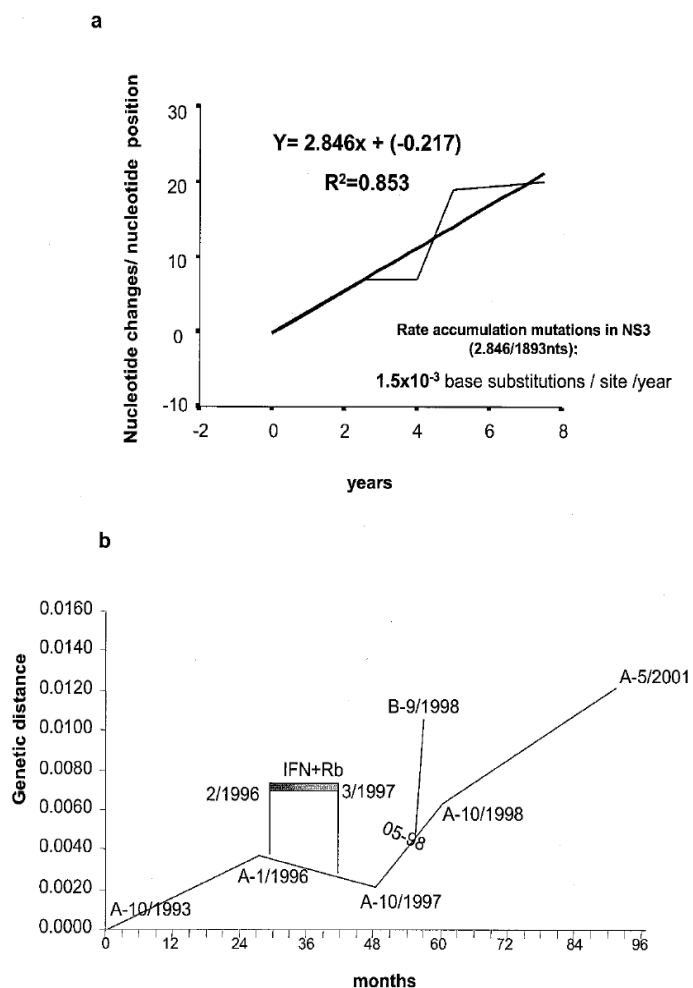


FIG. 3. (a) Nucleotide mutations per nucleotide position for each consecutive sample plotted against the year the sample was collected and used for regression analysis. (b) Representation of genetic distances between consensus nucleotide sequences from consecutive samples in patient A and an acute-phase sample from patient B. IFN+Rb, interferon plus Ribavirin.

tutions over time, and once divided by the number of nucleotides analyzed (1,893), it represents the rate of fixation of mutations per site and per year (35).

Genetic distance was calculated at the nucleotide and amino acid levels using the DNASP program following Jukes and Cantor algorithms which average the number of nucleotide substitutions per site between populations, considering each sample as a population (53).

HCV population parameters. The complexity of viral quasispecies (a quantitative estimate of the genetic information contained in the viral population) was evaluated according to two parameters: mutation frequency (Mf) and the ratio of nonsynonymous (dn) to synonymous (ds) mutations (dn/ds ratio). Mf was calculated as the total number of nucleotide or amino acid mutations relative to the consensus sequence of each sample divided by the total number of nucleotides or amino acids sequenced. Heterogeneity of the quasispecies increases as Mf increases (11). The dn/ds ratio (where dn is the frequency of nonsynonymous substitutions per nonsynonymous site and ds is the frequency of synonymous substitutions per synonymous site) was used as an estimate of differences in the type of pressure exerted on the evolving quasispecies (34, 36).

Quantitation of nucleotide misincorporation with *Pfu* DNA polymerase. To minimize misincorporation errors during PCR amplification, we used the high-fidelity DNA polymerase *Pfu* (Stratagene, San Diego, CA), which, under the conditions used, has a reported error rate of between 1.3×10^{-6} (8) and $6.5 \times$

10^{-7} (2). After 70 cycles of amplification, the expected number of mutations in the worst reported case is one error in 9.1×10^5 nucleotides sequenced. In our study, after sequencing 45,432 nucleotides, the expected number of errors due to *Pfu* is 0.5. In addition, a control study has been carried out, in which one of the clones (1,893 nucleotides) of known sequence was subjected to 70 additional cycles of PCR in 20 independent amplifications. No mutations were detected after analyzing 8,140 nucleotides. This result improves our previous estimates of the misincorporation rate when using a more error-prone polymerase, *Taq* DNA polymerase (Perkin-Elmer) (1/5,451 and 2/9,964) (4, 37), and confirms that most of the observed heterogeneity was independent of artifacts during the amplification procedure.

Statistical analysis. We compared population parameters of viral quasispecies from patient A with those of quasispecies from patient B. Normal distribution of quantitative variables was assessed by use of the Kolmogorov-Smirnoff test. Continuous variables were expressed as mean and standard deviation or median and interquartile range and compared by use of the *t* test or Mann-Whitney U test. Categorical variables were compared by the chi-square test with Yates correction or Fisher's exact test. All *P* values were two-sided, and differences were considered statistically significant when the *P* value was <0.05 . Linear regression and data analysis were performed using SPSS version 11.0 for Windows.

TABLE 1. Population parameters of quasiespecies of chronic-phase sample A-10/1998 from patient A and acute-phase sample B-9/1998 from patient B

Patient	Nucleotides		Amino acids	
	Proportion of sequences	No. of mutations	Proportion of sequences	No. of mutations
A	1/12	22	1/12	4
	1/12	21	1/12	3
	1/12	19	1/12	2
	1/12	12	7/12	1
	1/12	7	2/12	0
	2/12	5		
	1/12	4		
	4/12	3		
Median (interquartile range) or mean $\pm$ SE		9.5 (4.5–20)		2 $\pm$ 0.71
B	1/12	7	2/12	3
	1/12	5	2/12	2
	1/12	4	5/12	1
	3/12	3	3/12	0
	3/12	2		
	3/12	0		
Median (interquartile range) or mean $\pm$ SE		3.5 (2–5)		1.5 $\pm$ 0.65
P value		0.043		0.626

Nucleotide sequence accession numbers. The 30 nucleotide sequences generated in the present study are identified at GenBank with accession numbers DQ068178 to DQ068207.

RESULTS

Evolution of the NS3 consensus sequence in patient A over 8 years. Analysis of consecutive consensus nucleotide sequences in patient A revealed that 7 mutations appeared and were fixed, while 13 substitutions were only temporarily predominant (data not shown). One mutation (C4782T), after being predominant in 1996, disappeared from subsequent consensus sequences but was still present in some clones in the A-10/1998 sample, indicating its maintenance in the evolving quasiespecies as a minority component. Similarly, comparison of the deduced amino acid sequences evidenced changes that were temporarily predominant and then either negatively selected or overtaken by other variants, as well as amino acid substitutions which appeared and became fixed in the consensus of all subsequent samples.

As shown in Fig. 3a, the rate of fixation of mutations in the NS3 gene based on regression analysis of sequential samples obtained over the entire 8-year interval was 1.5×10^{-3} mutations/site/year, similar to that previously reported for NS3 when calculation were done by comparing two time points (1.33×10^{-3}) (41). The intersample genetic distance of the consensus sequences is shown in Fig. 3b. A 1.7-year interval of relative stasis coinciding with the period of antiviral treatment, during which the viral load decreased by about 1 log unit (from 2.8×10^6 to 1.7×10^5 IU/ml), can be seen, followed by a

marked increase in genetic distance between the two consecutive samples obtained after treatment discontinuation. As expected, the highest intersample genetic distance corresponds to the sequence obtained in the newly infected host (sample B-9/1998).

Comparison of viral quasiespecies in the chronic carrier and infected sexual partner. In patient A, 107 nucleotide mutations were found in the twelve clones from sample A-10/1998 (Table 1). The maximum nucleotide difference between clones was 10%, and no master sequence was identified. Only 16 (15%) of the 107 nucleotide substitutions were nonsynonymous. At the amino acid level, a master sequence represented by two identical clones was present. As shown in Fig. 4, of the 16 amino acid changes, 7 occurred within or flanking known CD4 or CD8 T-cell epitopes (9, 10, 62), 8 occurred in unknown functional regions, and 1 drastic mutation (A1349D) occurred in the well-conserved TATPP switch region of the helicase domain (65).

In patient B, a master sequence was detected at the nucleotide and amino acid levels, representing 25% of the clones which were identical to the consensus sequence (Table 1 and Fig. 4). Altogether, 31 nucleotide mutations were found within the 12 clones, with a maximum 3.7% difference between clones. In this patient, 15 (48%) of the nucleotide substitutions were nonsynonymous and most amino acid changes (67%) occurred in the proteinase domain. Nine changes occurred within known or flanking CD4 or CD8 T-cell epitopes (9, 10, 62), five occurred in unknown functional regions, and one (R1490K) occurred in a very well conserved residue of the nucleotide triphosphate binding motif VI, which has been reported to be essential for helicase activity (49).

The complexity of the viral quasiespecies at the nucleotide level was significantly greater in the chronic carrier than in the partner with acute resolving hepatitis, when considering both the median number of mutations per sequence and the overall mutation frequency (Table 1). The difference was not significant at the amino acid level. Reduction in genetic diversity in the quasiespecies of patient B was evident when synonymous sites were compared (ds), but no difference in viral diversity was apparent when nonsynonymous mutations (dn) were compared (Table 2).

Sixteen nucleotide substitutions (including the above-mentioned C4782T substitution) which were present in at least one of the clones of the A-10/1998 quasiespecies (clone A3) became predominant in the viral population of patient B (B-9/1998) but were not present in subsequent consensus sequences from patient A. As shown in Fig. 4, a signature pattern of three amino acid substitutions simultaneously present in the A3 clone of the A-10/1998 quasiespecies became predominant in the viral population of patient B. In addition, the consensus phylogenetic tree constructed with consecutive average nucleotide sequences from patient A and the 24 individual clones from both patients (Fig. 5) clearly showed that all sequences from patient B clustered with clone A3 from A-10/1998, suggesting a major role of this particular clone in the generation of the viral quasiespecies in patient B.

Despite the low error rate of the *Pfu* DNA polymerase, we cannot rule out the possibility that some individual mutations identified in the present study were in fact misincorporation

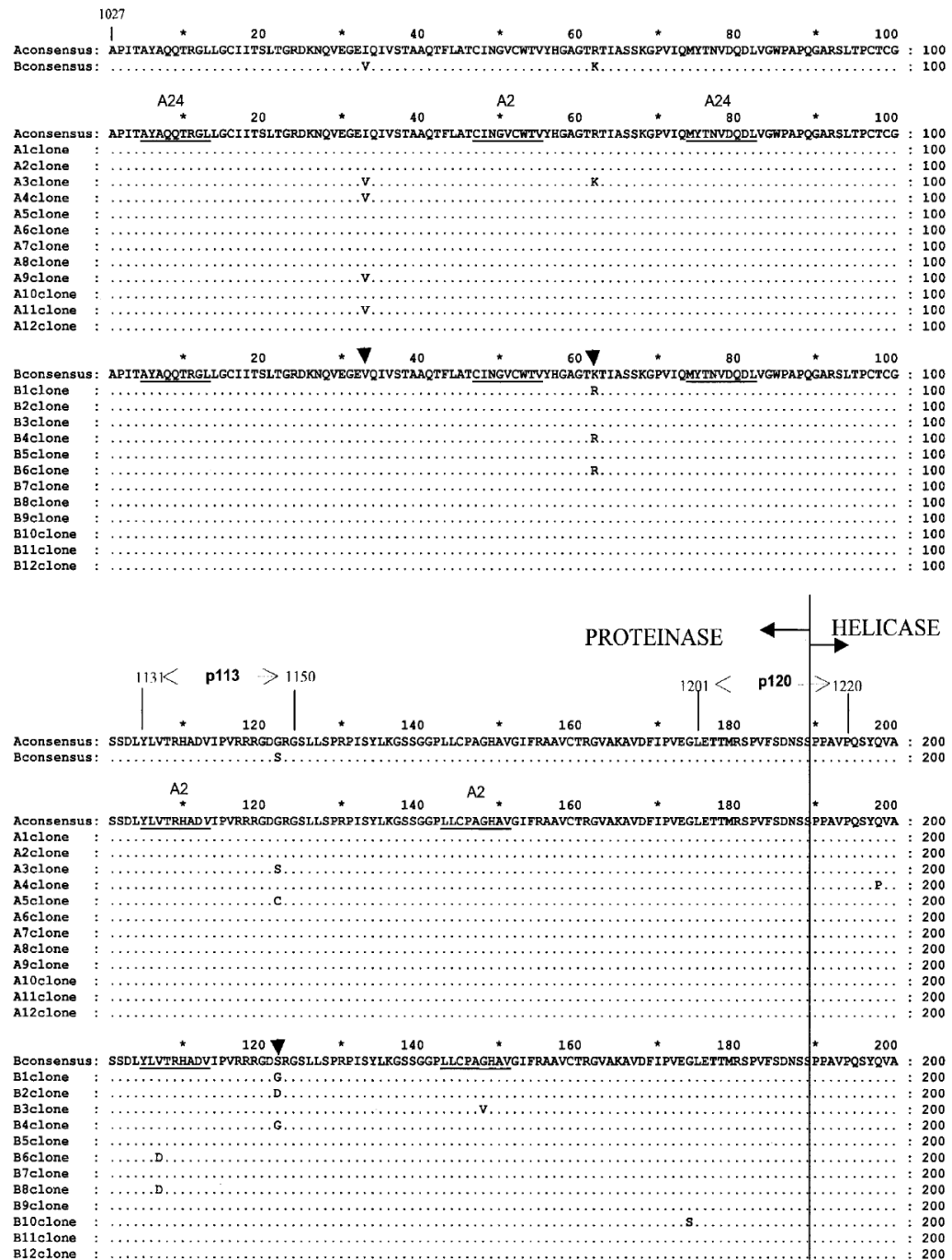


FIG. 4. Sequence alignment of the complete 631 amino acids of NS3 in 12 clones of patient A close to the time of transmission to patient B (sample A-10/1998, clones A1 to A12) and in 12 clones from patient B (B-9/1998, clones B1 to B12). Dots indicate residues identical to those of the consensus sequences. Arrowheads indicate amino acid changes in the consensus sequences of patients A and B. Reported HLA class I A-restricted epitopes are underlined, and B-restricted epitopes are shown in italics. HLA class II-restricted CD4 epitopes are indicated between vertical lines and labeled according to published reports (9, 10, 62).

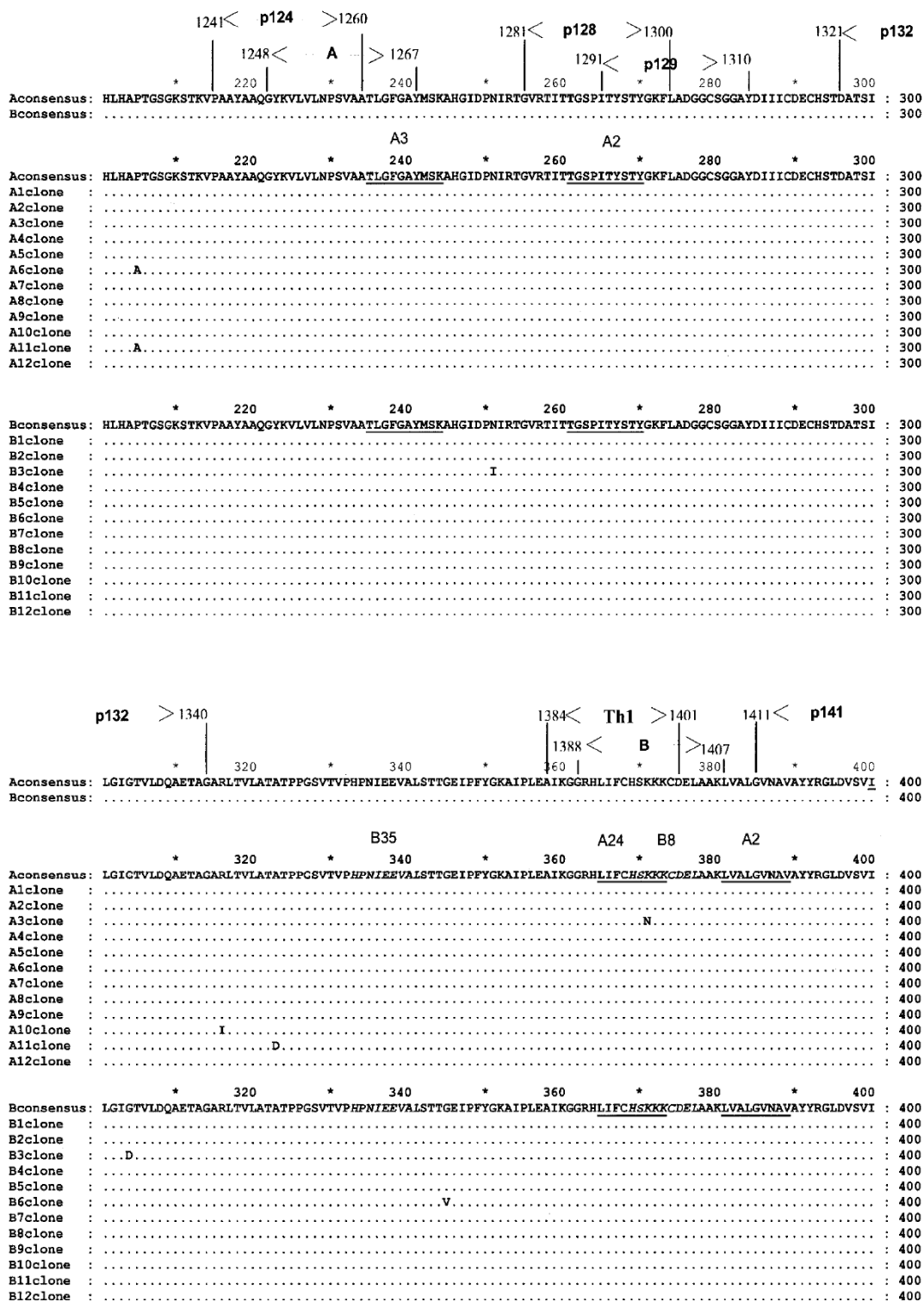


FIG. 4—Continued.

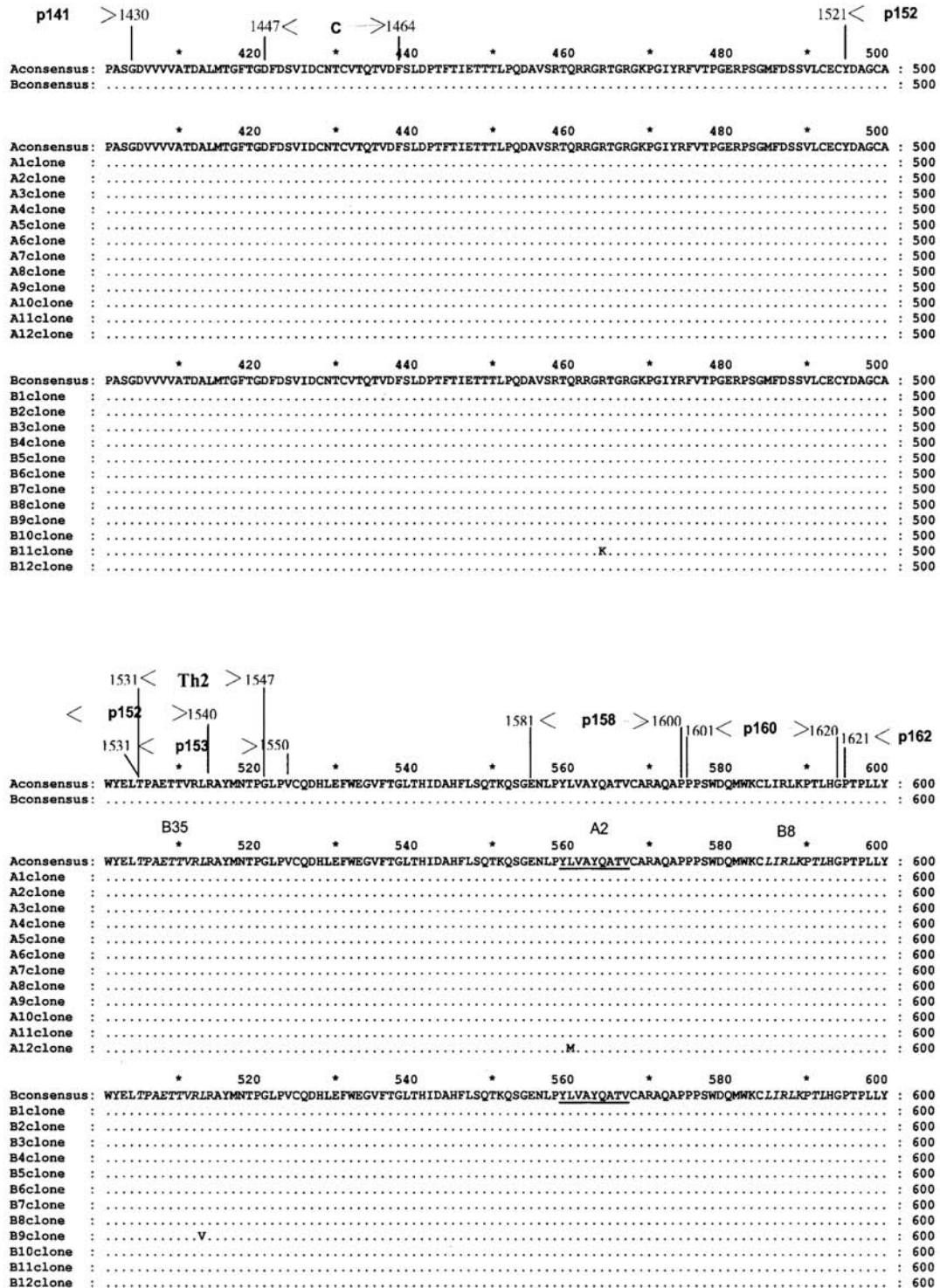


FIG. 4—Continued.

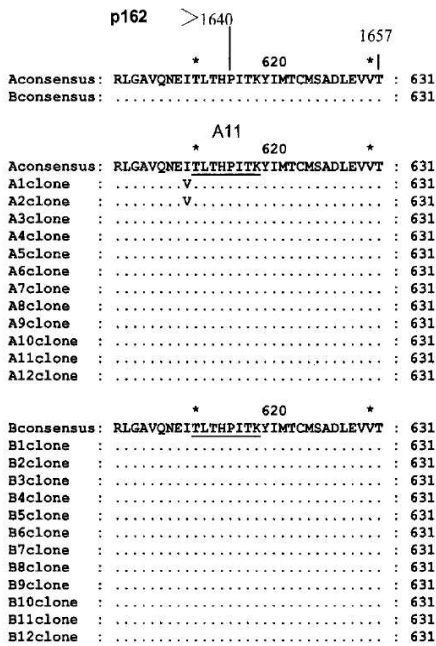


FIG. 4—Continued.

errors introduced during the reverse transcription step of the RT-PCR.

DISCUSSION

Determinants of viral clearance or persistence after HCV exposure are not well defined but are likely to include both viral and host factors. Among the former, it has been previously reported that genetic complexity is significantly higher in transfusion recipients than in intravenous drug users, suggesting that inoculum size influences emergence and development of the new viral quasispecies (42). Among the latter, both innate and adaptive immune responses are essential to keep viral replication under control and eventually clear infection. Transmission of HCV into a new host provides a unique opportunity to identify virus-related features associated with outcome of infection. Recent reports have provided evidence of positive Darwinian pressure leading to evolutionary changes within T-cell epitopes recognized by the host as well as reversion of prior escape mutants unrecognized by the new host (25, 30, 60, 63). Both selection and reversion of T-cell epitope

TABLE 2. Comparison of parameters of viral quasispecies complexity between patients A and B^a

Patient	Mutation frequency		Genetic divergence		
	Nucleotides	Amino acids	ds	dn	dn/ds
A	0.0047	0.0021	0.02681	0.00174	0.065
B	0.0021	0.0020	0.00468	0.00176	0.376
P value	<0.001	1	<0.001	0.831	<0.001

^a For abbreviations and mutation frequency units, see Materials and Methods.
^b P value refers to the comparison between patients A and B.

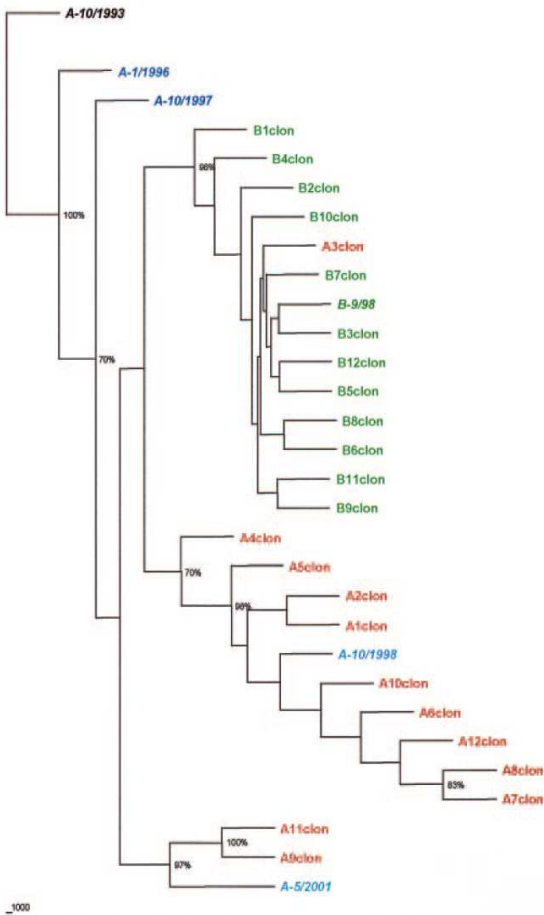


FIG. 5. Consensus phylogenetic tree showing frequencies of >70% in bootstrap analysis. Branch lengths are drawn to scale.

escape variants are influenced by their relative fitness cost to the replicating quasispecies (1).

In the current study we evaluated evolutionary changes in a relatively well-conserved genomic region of HCV encoding essential enzymatic functions and known to contain a large number of T-cell epitopes. As expected, during persistent infection, the rate of fixation of mutations was relatively stable despite the marked heterogeneity of the viral quasispecies. The finding that mutation C4782T, which was predominant in 1996 but not in later samples, was still maintained as a minority component in some clones in the A-10/1998 sample confirms a feature of the dynamics of viral quasispecies, namely, that because of perturbations in environmental conditions, both master sequences and any given mutant distribution often have a fleeting existence. It may also represent an example of the molecular memory proposal in evolving RNA viral quasispecies (54) in the form of genomes hidden in the mutant spectrum as a molecular record of the past evolutionary history of the population. Molecular memory of viral quasispecies has been documented for several genetic markers in several viruses (16, 54–56). The presence of memory genomes in the mutant

spectrum of the evolving quasispecies is likely to represent an advantage by providing a mechanism to respond efficiently to a selective pressure that has been previously experienced by the viral population.

When transmitted into a new host, a decrease in viral quasispecies complexity was observed along with an increase in the proportion of nonsynonymous mutations. These findings should not be surprising during acute infection of a host with a different HLA haplotype. Because of the cross-sectional nature of the data, the significantly higher proportion of nonsynonymous mutations in the viral quasispecies of patient B cannot be interpreted as the result of positive selection and may only suggest differences in the type of selection exerted on the viral population in this patient. On the other hand, the excess of synonymous mutations in the viral quasispecies in patient A cannot be attributed to purifying selection, since in RNA viruses the RNA itself is an important part of the viral phenotype and silent mutations may also be subject to selection (40). Unexpectedly, however, the consensus sequence of the entire NS3 gene at the nucleotide and amino acid level was almost identical to that of a clone (A3) present in the infecting quasispecies, despite the complete disparity in HLA class I and II haplotypes between the patients and evidence of a strong immune response to the protein in patient B as assessed by gamma interferon enzyme-linked immunospot assay (data not shown). Since viral load was not different, the extremely low level of fixation of mutations after 7 weeks of infection can be explained only by the effect of genetic bottlenecks. In this regard, the drastic reduction in the number and diversity of infectious particles starting infection, along with a reduced adaptability of the randomly sampled genomes carrying potentially deleterious mutations, might have restricted the generation of mutants able to overcome the selective pressure exerted by the innate and adaptive immune responses in the new host (39, 46, 47). Effects of genetic bottlenecks are stochastic and hence unpredictable, and although bottlenecks may tend to favor spontaneous resolution, persistent infection may also ensue.

Genetic bottlenecks might operate in circumstances in which HCV transmission occurs through exposure to a very small inoculum. Male-to-female sexual transmission may be a good example, since HCV RNA has been detected at very low levels in semen (33, 64). Although sexual transmission of HCV has generally been considered rare because of the low prevalence of anti-HCV among sexual partners of long-term carriers, its frequency may be much more common than originally thought. Evidence of HCV-specific T-cell responses directed to nonstructural antigens in the absence of other markers of exposure has been found in over 40% of sexual contacts of chronic carriers in retrospective studies (3) and in more than 60% of prospectively followed contacts of acute hepatitis patients (29). In the case described here, the effective inoculum size might have been relatively larger than usual because of vaginal mucosal damage during removal of an intrauterine device, and the sample of sequences starting infection were fit enough to overcome innate immune responses and reach high-level replication. A strong adaptive immune response by the HLA-mismatched host subsequently led to acute symptomatic hepatitis and eradication of a viral population unable to generate sufficient diversity to overcome immune-mediated clearance.

In summary, our results suggest that genetic bottlenecks during sexual transmission of HCV, with its associated reduction in viral diversity and potential presence of restrictive mutations (6, 12, 38, 46), may help explain the higher likelihood of spontaneous recovery after low-level HCV exposure and provide another example of how the quasispecies structure of most RNA viruses may influence infection outcome (15).

ACKNOWLEDGMENTS

This work was supported by grants from the Fundació Privada Catalana d'Hemofilia, Barcelona, Spain; the Fondo de Investigaciones Sanitarias (grant PI030274); the Spanish Ministerio de Educación y Ciencia (grant SAF2003-08724); the Red Nacional de Investigación en Hepatología y Gastroenterología (RNIHG) (grant C03/02); and the Fundació Marató TV3 (grants 00/2010 and 00/2110).

We are indebted to Esteban Domingo for helpful comments.

REFERENCES

- Altman, J. D., and M. B. Feinberg. 2004. HIV escape: there and back again. *Nat. Med.* 10:229–230.
- Andre, P., A. Kim, K. Khrapko, and W. G. Thilly. 1997. Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. *Genome Res.* 7:843–852.
- Bronowicki, J. P., D. Vetter, G. Uhl, H. Hudziak, A. Uhrmacher, J. M. Vetter, and M. Doffoel. 1997. Lymphocyte reactivity to hepatitis C virus (HCV) antigens shows evidence for exposure to HCV in HCV-seronegative spouses of HCV-infected patients. *J. Infect. Dis.* 176:518–522.
- Cabot, B., M. Martell, J. I. Esteban, M. Piron, T. Otero, R. Esteban, J. Guardia, and J. Gomez. 2001. Longitudinal evaluation of the structure of replicating and circulating hepatitis C virus quasispecies in nonprogressive chronic hepatitis C patients. *J. Virol.* 75:12005–12013.
- Cerny, A., and F. V. Chisari. 1999. Pathogenesis of chronic hepatitis C: immunological features of hepatic injury and viral persistence. *Hepatology* 30:595–601.
- Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 348:454–455.
- Clarke, D. K., E. A. Duarte, A. Moya, S. F. Elena, E. Domingo, and J. Holland. 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *J. Virol.* 67:222–228.
- Cline, J., J. C. Braman, and H. H. Hogrefe. 1996. PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* 24:3546–3551.
- Day, C. L., G. M. Lauer, G. K. Robbins, B. McGovern, A. G. Wurcel, R. T. Gandhi, R. T. Chung, and B. D. Walker. 2002. Broad specificity of virus-specific CD4⁺ T-helper-cell responses in resolved hepatitis C virus infection. *J. Virol.* 76:12584–12595.
- Diepolder, H. M., J. T. Gerlach, R. Zachoval, R. M. Hoffmann, M. C. Jung, E. A. Wierenga, S. Scholz, T. Santantonio, M. Houghton, S. Southwood, A. Sette, and G. R. Pape. 1997. Immunodominant CD4⁺ T-cell epitope within nonstructural protein 3 in acute hepatitis C virus infection. *J. Virol.* 71:6011–6019.
- Domingo, E. 1996. Biological significance of viral quasispecies. *Viral Hep. Rev.* 2:247–261.
- Domingo, E., C. Escarmis, N. Sevilla, A. Moya, S. F. Elena, J. Quer, I. S. Novella, and J. J. Holland. 1996. Basic concepts in RNA virus evolution. *FASEB J.* 10:859–864.
- Domingo, E., and J. J. Holland. 1988. High error rates, population equilibrium, and evolution of RNA replication systems, p. 3–36. *In* E. Domingo and J. J. Holland (ed.), *RNA genetics*. CRC Press, Boca Raton, Fla.
- Domingo, E., and J. J. Holland. 1994. Mutation rates and rapid evolution of RNA viruses, p. 161–184. *In* S. S. Morse (ed.), *Evolutionary biology of viruses*. Raven Press, New York, N.Y.
- Domingo, E., and J. J. Holland. 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* 51:151–178.
- Domingo, E., C. M. Ruiz-Jarabo, S. Sierra, A. Arias, N. Pariente, E. Baranowski, and C. Escarmis. 2002. Emergence and selection of RNA virus variants: memory and extinction. *Virus Res.* 82:39–44.
- Domingo, E., D. Sabo, T. Taniguchi, and C. Weissmann. 1978. Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13:735–744.
- Duarte, E., D. Clarke, A. Moya, E. Domingo, and J. Holland. 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc. Natl. Acad. Sci. USA* 89:6015–6019.
- Duarte, E. A., I. S. Novella, S. C. Weaver, E. Domingo, S. Wain-Hobson, D. K. Clarke, A. Moya, S. F. Elena, J. C. De La Torre, and J. J. Holland. 1994. RNA virus quasispecies: significance for viral disease and epidemiology. *Infect. Agents Dis.* 3:201–214.
- Eigen, M., and C. K. Biebricher. 1988. Variability of RNA genomes. *Se-*

- quence space and quasispecies distribution, p. 211–245. In E. Domingo, J. J. Holland, and P. Ahlquist (ed.), *RNA genetics*. CRC Press, Inc., Boca Raton, Fla.
21. Eigen, M., and P. Schuster. 1979. The hypercycle. A principle of natural self-organization. Springer-Verlag, Berlin, Germany.
 22. Escarmis, C., M. Davila, N. Charpentier, A. Bracho, A. Moya, and E. Domingo. 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. *J. Mol. Biol.* 264:255–267.
 23. Felsenstein, J. 2002. PHYLIP. Department of Genetics, University of Washington, Seattle. <http://evolution.genetics.washington.edu/phylip.html>.
 24. Genetic Computer Group. 2002. CLUSTALW. Wisconsin package version 8.1. Genetic Computer Group, Madison, Wis.
 25. Grakoui, A., N. H. Shoukry, D. J. Woollard, J. H. Han, H. L. Hanson, J. Ghayeb, K. K. Murthy, C. M. Rice, and C. M. Walker. 2003. HCV persistence and immune evasion in the absence of memory T cell help. *Science* 302:659–662.
 26. Grande Gimenez, M. C., F. El Far, W. S. Barsanti, and E. A. Servolo Medeiros. 2001. Cut and puncture accidents involving health care workers exposed to biological materials. *Braz. J. Infect. Dis.* 5:235–242.
 27. Holland, J. J., J. C. De La Torre, and D. A. Steinhauer. 1992. RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* 176:1–20.
 28. Holland, J. J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol, and S. VandePol. 1982. Rapid evolution of RNA genomes. *Science* 215:1577–1585.
 29. Kamal, S. M., A. Amin, M. Madwar, C. S. Graham, Q. He, A. Al Tawil, J. Rasenack, T. Nakano, B. Robertson, A. Ismail, and M. J. Koziel. 2004. Cellular immune responses in seronegative sexual contacts of acute hepatitis C patients. *J. Virol.* 78:12252–12258.
 30. Karlsson, A. C., H. Gaines, M. Sallberg, S. Lindback, and A. Sonnerborg. 1999. Reappearance of founder virus sequence in human immunodeficiency virus type 1-infected patients. *J. Virol.* 73:6191–6196.
 31. Koziel, M. J., D. K. Wong, D. Dudley, M. Houghton, and B. D. Walker. 1997. Hepatitis C virus-specific cytolytic T lymphocyte and T helper cell responses in seronegative persons. *J. Infect. Dis.* 176:859–866.
 32. Kwok, S., and R. Higuchi. 1989. Avoiding false positives with PCR. *Nature* 339:237–238.
 33. Leruez-Ville, M., J. M. Kunstmann, M. De Almeida, C. Rouzioux, and M. L. Chaix. 2000. Detection of hepatitis C virus in the semen of infected men. *Lancet* 356:42–43.
 34. Li, W. H., C. I. Wu, and C. C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* 2:150–174.
 35. Lu, L., T. Nakano, E. Orito, M. Mizokami, and B. H. Robertson. 2001. Evaluation of accumulation of hepatitis C virus mutations in a chronically infected chimpanzee: comparison of the core, E1, HVR1, and NS5b regions. *J. Virol.* 75:3004–3009.
 36. Manzin, A., L. Solfrosi, M. Debiaggi, F. Zara, E. Tanzi, L. Romano, A. R. Zanetti, and M. Clementi. 2000. Dominant role of host selective pressure in driving hepatitis C virus evolution in perinatal infection. *J. Virol.* 74:4327–4334.
 37. Martell, M., J. I. Esteban, J. Quer, J. Genesca, A. Weiner, R. Esteban, J. Guardia, and J. Gomez. 1992. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J. Virol.* 66:3225–3229.
 38. Muller, H. J. 1964. The relation of recombination to mutational advance. *Mutat. Res.* 106:2–9.
 39. Novella, I. S. 2004. Negative effect of genetic bottlenecks on the adaptability of vesicular stomatitis virus. *J. Mol. Biol.* 336:61–67.
 40. Novella, I. S., S. Zarate, D. Metzgar, and Ebendick-Corp. 2004. Positive selection of synonymous mutations in vesicular stomatitis virus. *J. Mol. Biol.* 342:1415–1421.
 41. Okamoto, H., M. Kojima, S. Okada, H. Yoshizawa, H. Iizuka, T. Tanaka, E. E. Muchmore, D. A. Peterson, Y. Ito, and S. Mishiro. 1992. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. *Virology* 190:894–899.
 42. Pawlotsky, J. M., M. Pellerin, M. Bouvier, F. Roudot-Thoraval, G. Germanidis, A. Bastie, F. Darthuy, J. Remire, C. J. Soussy, and D. Dhumeaux. 1998. Genetic complexity of the hypervariable region 1 (HVR1) of hepatitis C virus (HCV): influence on the characteristics of the infection and responses to interferon alpha therapy in patients with chronic hepatitis C. *J. Med. Virol.* 54:256–264.
 43. Penin, F., J. Dubuisson, F. A. Rey, D. Moradpour, and J. M. Pawlotsky. 2004. Structural biology of hepatitis C virus. *Hepatology* 39:5–19.
 44. Puro, V., N. Petrosillo, G. Ippolito, et al. 1995. Risk of hepatitis C seroconversion after occupational exposures in health care workers. *Am. J. Infect. Control* 23:273–277.
 45. Quer, J., and J. I. Esteban. 2005. Epidemiology, p. 407–425. In H. C. Thomas, S. Lemon, and A. J. Zuckerman (ed.), *Viral hepatitis*. Blackwell Publishing, Oxford, United Kingdom.
 46. Quer, J., C. L. Hershey, E. Domingo, J. J. Holland, and I. S. Novella. 2001. Contingent neutrality in competing viral populations. *J. Virol.* 75:7315–7320.
 47. Quer, J., R. Huerta, I. S. Novella, L. Tsimring, E. Domingo, and J. J. Holland. 1996. Reproducible nonlinear population dynamics and critical points during replicative competitions of RNA virus quasispecies. *J. Mol. Biol.* 264:465–471.
 48. Quer, J., P. Murillo, J. I. Esteban, M. Martell, R. Esteban, and J. Guardia. 2003. Sexual transmission of hepatitis C virus from a patient with chronic disease to his sex partner after removal of an intrauterine device. *Sex. Transm. Dis.* 30:470–471.
 49. Rho, J., S. Choi, Y. R. Seong, J. Choi, and D. S. Im. 2001. The arginine-1493 residue in QRRGRTGR1493G motif IV of the hepatitis C virus NS3 helicase domain is essential for NS3 protein methylation by the protein arginine methyltransferase 1. *J. Virol.* 75:8031–8044.
 50. Roderic, D. M. 1997. TREEVIEW v1.6.5. University of Glasgow, Scotland, United Kingdom. <http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>.
 51. Rosen, H. R., C. Miner, A. W. Sasaki, D. M. Lewinsohn, A. J. Conrad, A. Bakke, H. G. Bouwer, and D. J. Hinrichs. 2002. Frequencies of HCV-specific effector CD4+ T cells by flow cytometry: correlation with clinical disease stages. *Hepatology* 35:190–198.
 52. Ross, R. S., S. Viazov, M. Thormahlen, L. Bartz, J. Tamm, P. Rautenberg, M. Roggendorf, and A. Deister. 2002. Risk of hepatitis C virus transmission from an infected gynecologist to patients: results of a 7-year retrospective investigation. *Arch. Intern. Med.* 162:805–810.
 53. Rozas, J., and R. Rozas. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15:174–175.
 54. Ruiz-Jarabo, C. M., A. Arias, E. Baranowski, C. Escarmis, and E. Domingo. 2000. Memory in viral quasispecies. *J. Virol.* 74:3543–3547.
 55. Ruiz-Jarabo, C. M., A. Arias, C. Molina-Paris, C. Briones, E. Baranowski, C. Escarmis, and E. Domingo. 2002. Duration and fitness dependence of quasispecies memory. *J. Mol. Biol.* 315:285–296.
 56. Ruiz-Jarabo, C. M., E. Miller, G. Gomez-Mariano, and E. Domingo. 2003. Synchronous loss of quasispecies memory in parallel viral lineages: a deterministic feature of viral quasispecies. *J. Mol. Biol.* 333:553–563.
 57. Scognamiglio, P., D. Accapezzato, M. A. Casciaro, A. Cacciani, M. Artini, G. Bruno, M. L. Chircu, J. Sidney, S. Southwood, S. Abrignani, A. Sette, and V. Barnaba. 1999. Presence of effector CD8+ T cells in hepatitis C virus-exposed healthy seronegative donors. *J. Immunol.* 162:6681–6689.
 58. Shata, M. T., N. Tricoche, M. Perkus, D. Tom, B. Brotman, P. McCormack, W. Pfahler, D. H. Lee, L. H. Tobler, M. Busch, and A. M. Prince. 2003. Exposure to low infective doses of HCV induces cellular immune responses without consistently detectable viremia or seroconversion in chimpanzees. *Virology* 314:601–616.
 59. Steinhauer, D. A., and J. J. Holland. 1987. Rapid evolution of RNA viruses. *Annu. Rev. Microbiol.* 41:409–433.
 60. Tester, L., S. Smyk-Pearson, P. Wang, A. Wertheimer, E. Yao, D. M. Lewinsohn, J. E. Tavis, and H. R. Rosen. 2005. Immune evasion versus recovery after acute hepatitis C virus infection from a shared source. *J. Exp. Med.* 201:1725–1731.
 61. Villano, S. A., D. Vlahov, K. E. Nelson, S. Cohn, and D. L. Thomas. 1999. Persistence of viremia and the importance of long-term follow-up after acute hepatitis C infection. *Hepatology* 29:908–914.
 62. Wang, H., and D. D. Eckels. 1999. Mutations in immunodominant T cell epitopes derived from the nonstructural 3 protein of hepatitis C virus have the potential for generating escape variants that may have important consequences for T cell recognition. *J. Immunol.* 162:4177–4183.
 63. Weiner, A., A. L. Erickson, J. Kansopon, K. Crawford, E. Muchmore, A. L. Hughes, M. Houghton, and C. M. Walker. 1995. Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escape variant. *Proc. Natl. Acad. Sci. USA* 92:2755–2759.
 64. Wejstal, R. 1999. Sexual transmission of hepatitis C virus. *J. Hepatol.* 31(Suppl. 1):92–95.
 65. Yao, N., T. Hesson, M. Cable, Z. Hong, A. D. Kwong, H. V. Le, and P. C. Weber. 1997. Structure of the hepatitis C virus RNA helicase domain. *Nat. Struct. Biol.* 4:463–467.
 66. Zuckerman, J., G. Clewley, P. Griffiths, and A. Cockcroft. 1994. Prevalence of hepatitis C antibodies in clinical health-care workers. *Lancet* 343:1618–1620.

ESTUDIO 2

4 ESTUDIO 2: "NATURALLY OCURRING NS3-PROTEASE-INHIBITOR RESISTANT MUTANT A156T IN THE LIVER OF AN UNTREATED CHRONIC HEPATITIS C PATIENT"

4.1 Hipótesis

Las mutaciones de resistencia a antivirales del VHC son generadas constantemente dentro de la quasiespecie de todos los pacientes con infección crónica por VHC de manera natural. Estos mutantes aparecen como componentes minoritarios en la población viral de un paciente crónico antes del tratamiento. Se estudiará la presencia de estos mutantes mediante clonación-secuenciación por Sanger.

4.2 Objetivo

Estudio de la presencia de mutantes de resistencia a inhibidores de la proteasa NS3 en un paciente VHC crónico no tratado con estos antivirales.

4.3 Resumen

El tratamiento actual de la infección por VHC consiste en la combinación de ribavirina con interferón pegilado, que conduce a una respuesta virológica sostenida en aproximadamente el 30-45% de los pacientes infectados por VHC genotipo 1. Dada la eficacia limitada del mismo y el desarrollo frecuente de efectos adversos, se están desarrollando tratamientos más eficaces. Más de 50 nuevos antivirales (inhibidores de la proteasa NS3, NS5A y de la ARN-polimerasa NS5B), globalmente designados DAAs ("Direct antiviral agents"), e inhibidores de proteínas celulares esenciales en el ciclo biológico viral (ciclofilina B), están actualmente en

diferentes fases de estudios clínicos, algunos en fase III, que combinados con el tratamiento estándar, han reportado resultados prometedores en pacientes con VHC genotipo 1. La causa principal de fallo al tratamiento con estos inhibidores es la aparición de mutaciones de resistencia.

A partir del conocimiento de la enorme variabilidad que presenta el VHC, el objetivo de nuestro trabajo fue buscar mediante clonación y secuenciación por Sanger, si estos mutantes estaban presentes en la población viral de un aislado de hígado de un paciente crónico.

A partir de una biopsia hepática de un paciente no tratado con inhibidores de proteasa, se obtuvieron un total de 128 clones del dominio NS3-proteasa (567 nucleótidos) mediante la técnica de clonación-secuenciación por Sanger. El estudio confirmó que es una región muy conservada, donde el 86% (110/128) de las secuencias fueron iguales, representando una secuencia máster que coincidió con la secuencia consenso. El resultado principal fue la detección de una mutación, el cambio de una Alanina por una Treonina en la posición 156 (A156T) en la quasiespecies viral del paciente A, con una frecuencia cercana al 1%. La mutación A156T se ha descrito en la bibliografía que confiere niveles elevados de resistencia a los inhibidores de NS3: telaprevir y boceprevir (aprobados para su uso en clínica).

La presencia de mutaciones que confieren resistencia a antivirales, en un paciente, no tratado es una consecuencia directa de la continua generación de variantes, y este hecho representa un inconveniente importante en el tratamiento de las enfermedades virales.

Estudios diferentes de secuenciación directa han puesto de manifiesto variantes resistentes como cepas dominantes preexistentes en las posiciones V36, T54, R155, D168 Y V170 de la proteasa en el 0,2-2,2% de los pacientes (108-111). La aparición rápida de variantes resistentes durante el tratamiento con agentes antivirales en monoterapia sugiere que estas cepas virales pre-existen en la población viral. Esto sugiere la necesidad de caracterizar el perfil de resistencia a los agentes antivirales del VHC con el fin de aumentar la comprensión de los mecanismos de mutaciones de resistencia.

Dada la limitación de la clonación y secuenciación por Sanger y la aparición de las técnicas de secuenciación de alto rendimiento, se analizó la población viral de este paciente con una mayor profundidad y de otros aislados de pacientes con infección por VHC en el estudio 3.



Available online at www.sciencedirect.com



Virology 370 (2008) 237–245

VIROLOGY

www.elsevier.com/locate/yviro

Rapid Communication

Naturally occurring NS3-protease-inhibitor resistant mutant A156T in the liver of an untreated chronic hepatitis C patient

Maria Cubero^{a,b,c}, Juan Ignacio Esteban^{a,b,c}, Teresa Otero^a, Silvia Sauleda^{c,d}, Marta Bes^{c,d}, Rafael Esteban^{a,b,c}, Jaume Guardia^{a,b,c}, Josep Quer^{a,b,c,*}

^a Liver Unit, Department of Medicine, Hospital Universitari Vall d'Hebron, Barcelona, Spain

^b Universitat Autònoma de Barcelona, UAB Bellaterra, Spain

^c CIBER de Enfermedades Hepáticas y Digestivas (CIBERehd) del Instituto de Salud Carlos III, Madrid, Spain

^d Banc de Sang i de Teixits, Institut Català de la Salut, Barcelona, Spain

Received 7 September 2007; returned to author for revision 17 September 2007; accepted 5 October 2007

Abstract

An increasing number of new hepatitis C virus NS3-protease inhibitors are being evaluated for the treatment of chronic hepatitis C. Treatment-induced selection of mutants conferring resistance to protease inhibitors has been shown both in vivo and in vitro. A specific mutation, A156T has been shown to confer high-level resistance to several such agents (BILN2061, VX-950, SCH446211 (SCH6) and SCH503034). Here we report the presence of the A156T mutation in close to 1% of NS3 sequences within the liver quasispecies of a chronic hepatitis C patient never treated with anti-NS3-protease inhibitors.

© 2007 Elsevier Inc. All rights reserved.

Keywords: HCV-NS3; Inhibitor; Mutant; Resistant; Untreated; Patient

Introduction

Hepatitis C virus (HCV) infection, affecting 3% of the human population (210 million people worldwide), has become one of the leading causes of chronic hepatitis, end-stage liver disease, and hepatocellular carcinoma. (Quer and Esteban, 2005). Like most RNA viruses, HCV evolves rapidly due to high mutation rates and high-level viral replication through an error-prone RNA polymerase lacking proofreading capacity. Consequently, in the infected individual, the viral population is composed of a complex mixture of different but closely related genomes known as quasispecies (Martell et al., 1992) whose shape is subject to continuous changes due both, to competitive selection of continuously arising mutants (Domingo and Holland, 1988, 1994; Duarte et al., 1994; Eigen and Biebricher, 1988; Holland et al., 1982, 1992; Steinhauer and Holland, 1987) and to cooperation

of variants with different phenotypes and fitness (Vignuzzi et al., 2006). Among the numerous implications of the quasispecies structure on the biology of RNA viruses, is that it predicts the presence in any given patient of all possible genomic variants carrying single nucleotide substitutions at each position, conditioned to the relative fitness of each individual mutant. The presence in a single host of all single mutants and double or triple mutants depends on the viral load (with 4×10^4 particles all simple mutants would be represented, with 1.6×10^9 particles all double mutants and 6.4×10^{13} for all single, double and triple mutants). With high-level HCV replication in the untreated patient (10^{12} new viral particles per day), point mutations at each position in the genome arise at least once every day, so potentially resistant mutants are constantly being generated.

Currently, the most effective treatment of persistent HCV infection involves the combination of the nucleoside analogue ribavirin with pegylated interferon alpha, which leads to a sustained virological response in ~50–55% of patients (less than 40% of those infected with genotype 1 and over 80% of those infected with HCV genotypes 2 or 3) (Fried et al., 2002), (Manns et al., 2001). Because of the limited efficacy of this

* Corresponding author. Liver Unit, Department of Internal Medicine, Hospital Universitari Vall d'Hebron, Pg Vall d'Hebron 119-129, 08035 Barcelona, Spain. Fax: +34 934894032.

E-mail address: jquer@ir.vhebron.net (J. Quer).

		*	20	*	40	*	60	*	80	
CONSENSUS	:	GCGCCCATCACGGCGTACGCCAGCAGACGAGGGGCCTCTTGGGATGCATAATCACCAGCTTGACCGGCCGGGACAAAA	:	80						
MastLiv06	:		:	80						
MutLiv5	:		:	80						
MutLiv3	:		:	80						
MutLiv2	:		:	80						
AE4-2006	:		:	80						
AF11-2006	:		:	80						
AH4-2006	:	C.....	:	80						
A1b-2006	:	A.....C.T.....	:	80						
A2b-2006	:		:	80						
A3b-2006	:		:	80						
A6b-2006	:		:	80						
A8b-2006	:		:	80						
A10b-2006	:		:	80						
B2b-2006	:		:	80						
B3b-2006	:		:	80						
B4b-2006	:		:	80						
B6b-2006	:		:	80						
B9b-2006	:		:	80						
B12b-2006	:		:	80						
C5b-2006	:		:	80						
C8b-2006	:		:	80						
D10b-2006	:		:	80						
F8b-2006	:		:	80						
F11b-2006	:		:	80						
F12b-2006	:		:	80						
G9b-2006	:		:	80						
H10b-2006	:	T.....	:	80						
H12b-2006	:		:	80						
E6b-2006	:		:	80						
E10b-2006	:		:	80						
SA7-2006	:		:	80						
SC4-2006	:		:	80						

		*	100	*	120	*	140	*	160	
CONSENSUS	:	CCAGGTGGAGGGTGAGGTTTCAGATCGTGTCAACTGCTGCCAGACTTTCTTAGCAACCTGCATTAAACGGAGTGTGCTGGA	:	160						
MastLiv06	:		:	160						
MutLiv5	:		:	160						
MutLiv3	:		:	160						
MutLiv2	:		:	160						
AE4-2006	:		:	160						
AF11-2006	:		:	160						
AH4-2006	:		:	160						
A1b-2006	:		:	160						
A2b-2006	:		:	160						
A3b-2006	:		:	160						
A6b-2006	:	A.....	:	160						
A8b-2006	:	A.....	:	160						
A10b-2006	:	T.....	:	160						
B2b-2006	:		:	160						
B3b-2006	:	T.....	:	160						
B4b-2006	:		:	160						
B6b-2006	:		:	160						
B9b-2006	:		:	160						
B12b-2006	:	G.....	:	160						
C5b-2006	:		:	160						
C8b-2006	:		:	160						
D10b-2006	:		:	160						
F8b-2006	:		:	160						
F11b-2006	:		:	160						
F12b-2006	:		:	160						
G9b-2006	:		:	160						
H10b-2006	:		:	160						
H12b-2006	:		:	160						
E6b-2006	:		:	160						
E10b-2006	:		:	160						
SA7-2006	:		:	160						
SC4-2006	:		:	160						

Fig. 1. Alignment of the sequences that were different at the nucleotide level obtained from the liver of the protease-inhibitor untreated chronic HCV G1a patient. The percentage at which each sequence was present in the isolate is shown at the end of each sequence. Dots indicate residues identical to those of the consensus sequence. Mutations at position 156 that confer resistance to the protease inhibitors are highlighted (bold and italics) and marked in between two vertical lines that limit the codon.

		*	180	*	200	*	220	*	240	
CONSENSUS	:	CTGTCTAC	CGGGGCCG	GAACAAGG	ACCATCGC	GTCTCTA	AGGGTCCT	GTTCATCC	AGATGTAC	ACCAATGTAGACCAA : 240
MastLiv06	:									: 240
MutLiv5	:									: 240
MutLiv3	:			A.....						: 240
MutLiv2	:									: 240
AE4-2006	:									: 240
AF11-2006	:									: 240
AH4-2006	:									: 240
A1b-2006	:									: 240
A2b-2006	:									: 240
A3b-2006	:									: 240
A6b-2006	:									: 240
A8b-2006	:									: 240
A10b-2006	:									: 240
B2b-2006	:									: 240
B3B-2006	:			A.....						: 240
B4b-2006	:			G.....						: 240
B6b-2006	:									: 240
B9b-2006	:									: 240
B12b-2006	:									: 240
C5b-2006	:									: 240
C8b-2006	:									: 240
D10b-2006	:									: 240
F8b-2006	:									: 240
F11b-2006	:									: 240
F12b-2006	:									: 240
G9b-2006	:								C.....	: 240
H10b-2006	:									: 240
H12b-2006	:									: 240
E6b-2006	:			A.....						: 240
E10b-2006	:									: 240
SA7-2006	:			G.....						: 240
SC4-2006	:									: 240

		*	260	*	280	*	300	*	320	
CONSENSUS	:	GACCTCGT	AGGCTGG	CCCCGCT	CCCCAAGG	TGCCCGCT	CACTGACAC	CCCTGCACT	TGCGGCTC	CTCGGACCTTTACCTGGT : 320
MastLiv06	:									: 320
MutLiv5	:									: 320
MutLiv3	:									: 320
MutLiv2	:									: 320
AE4-2006	:									: 320
AF11-2006	:									: 320
AH4-2006	:									: 320
A1b-2006	:									: 320
A2b-2006	:									: 320
A3b-2006	:									: 320
A6b-2006	:									: 320
A8b-2006	:									: 320
A10b-2006	:									: 320
B2b-2006	:									: 320
B3B-2006	:									: 320
B4b-2006	:									: 320
B6b-2006	:		A.....							: 320
B9b-2006	:		A.....							: 320
B12b-2006	:		A.....							: 320
C5b-2006	:									: 320
C8b-2006	:									: 320
D10b-2006	:									: 320
F8b-2006	:								CC.....	: 320
F11b-2006	:									: 320
F12b-2006	:									: 320
G9b-2006	:		C.....							: 320
H10b-2006	:									: 320
H12b-2006	:		A.....				A.....			: 320
E6b-2006	:									: 320
E10b-2006	:									: 320
SA7-2006	:									: 320
SC4-2006	:									: 320

Fig. 1 (continued).

it, selective inhibitors of the enzymatic activity of non-al proteins of the virus (NS3 protease, and NS5 ase) have been developed.

Because the NS3 protease is essential for viral polyprotein processing and, hence, productive HCV replication (Kolykhalov et al., 2000), several specific inhibitors (VX-950, SCH6 and

		*	340	*	360	*	380	*	400	
CONSENSUS :	CACGAGGCACGCCGATGTCATTCCCGTACGCCGCGGGGTGATGGCAGGGGCAGCCTGCTTTCGCCCCGGGCCATCTCTT	:	400							
MastLiv06 :		:	400							
MutLiv5 :		:	400							
MutLiv3 :		:	400							
MutLiv2 :		:	400							
AE4-2006 :		:	400							
AF11-2006 :		:	400							
AH4-2006 :		:	400							
A1b-2006 :		:	400							
A2b-2006 :		:	400							
A3b-2006 :		:	400							
A6b-2006 :		:	400							
A8b-2006 :		:	400							
A10b-2006 :	A.....	:	400							
B2b-2006 :		:	400							
B3b-2006 :		:	400							
B4b-2006 :		:	400							
B6b-2006 :		:	400							
B9b-2006 :		:	400							
B12b-2006 :		:	400							
C5b-2006 :		:	400							
C8b-2006 :	T.....	:	400							
D10b-2006 :		:	400							
F8b-2006 :		:	400							
F11b-2006 :	A.....	:	400							
F12b-2006 :	A.....	:	400							
G9b-2006 :		:	400							
H10b-2006 :		:	400							
H12b-2006 :	A.....	:	400							
E6b-2006 :		:	400							
E10b-2006 :		:	400							
SA7-2006 :		:	400							
SC4-2006 :		:	400							

	*	420	*	440	*	460	*	480	
CONSENSUS :	ACCTGAAAAGGCTCCTCGGGGGGCCCACTGCTGTGCCCGCGGGACACGCCGTAGGCATTTTCAGATCCCGCGGTGTGCACC	:	480						
MastLiv06 :		:	480						
MutLiv5 :		:	480						
MutLiv3 :		:	480						
MutLiv2 :		:	480						
AE4-2006 :		:	480						
AF11-2006 :		:	480						
AH4-2006 :		:	480						
A1b-2006 :		:	480						
A2b-2006 :		:	480						
A3b-2006 :		:	480						
A6b-2006 :	A.....	:	480						
A8b-2006 :		:	480						
A10b-2006 :		:	480						
B2b-2006 :		:	480						
B3b-2006 :	A.....	:	480						
B4b-2006 :		:	480						
B6b-2006 :	A.....	:	480						
B9b-2006 :		:	480						
B12b-2006 :		:	480						
C5b-2006 :		:	480						
C8b-2006 :		:	480						
D10b-2006 :	G.....	:	480						
F8b-2006 :		:	480						
F11b-2006 :		:	480						
F12b-2006 :		:	480						
G9b-2006 :		:	480						
H10b-2006 :		:	480						
H12b-2006 :		:	480						
E6b-2006 :		:	480						
E10b-2006 :		:	480						
SA7-2006 :		:	480						
SC4-2006 :		:	480						

156	
GCC (A)	
GCA (A)	
ACC (T)	

Fig. 1 (continued).

		*	500	*	520	*	540	*	560	
CONSENSUS	:	CGTGGAGTGGCTAAAGCGGTGGACTTTATCCCCGTAGAGGCCTAGAGACAACCATGAGGTCCCGGTGTTCTCGGACAA	:	560						
MastLiv06	:		:	560						
MutLiv5	:	G.....	:	560						
MutLiv3	:		:	560						
MutLiv2	:	A.....	:	560						
AE4-2006	:	T.....	:	560						
AF11-2006	:	C.....	:	560						
AH4-2006	:		:	560						
A1b-2006	:		:	560						
A2b-2006	:		:	560						
A3b-2006	:	T.....	:	560						
A6b-2006	:		:	560						
A8b-2006	:		:	560						
A10b-2006	:		:	560						
B2b-2006	:		:	560						
B3b-2006	:		:	560						
B4b-2006	:		:	560						
B6b-2006	:		:	560						
B9b-2006	:		:	560						
B12b-2006	:		:	560						
C5b-2006	:	A.....	:	560						
C8b-2006	:		:	560						
D10b-2006	:		:	560						
F8b-2006	:		:	560						
F11b-2006	:		:	560						
F12b-2006	:		:	560						
G9b-2006	:		:	560						
H10b-2006	:		:	560						
H12b-2006	:		:	560						
E6b-2006	:		:	560						
E10b-2006	:		:	560						
SA7-2006	:	G.....C.....	:	560						
SC4-2006	:	C.....	:	560						

CONSENSUS	:	TTCCTCC	:	567	
MastLiv06	:		:	567	70.31%
MutLiv5	:		:	567	3.91%
MutLiv3	:		:	567	2.34%
MutLiv2	:		:	567	1.60%
AE4-2006	:		:	567	0.78%
AF11-2006	:		:	567	0.78%
AH4-2006	:		:	567	0.78%
A1b-2006	:		:	567	0.78%
A2b-2006	:		:	567	0.78%
A3b-2006	:		:	567	0.78%
A6b-2006	:		:	567	0.78%
A8b-2006	:		:	567	0.78%
A10b-2006	:		:	567	0.78%
B2b-2006	:		:	567	0.78%
B3b-2006	:		:	567	0.78%
B4b-2006	:		:	567	0.78%
B6b-2006	:		:	567	0.78%
B9b-2006	:		:	567	0.78%
B12b-2006	:		:	567	0.78%
C5b-2006	:		:	567	0.78%
C8b-2006	:		:	567	0.78%
D10b-2006	:		:	567	0.78%
F8b-2006	:		:	567	0.78%
F11b-2006	:		:	567	0.78%
F12b-2006	:		:	567	0.78%
G9b-2006	:		:	567	0.78%
H10b-2006	:		:	567	0.78%
H12b-2006	:		:	567	0.78%
E6b-2006	:		:	567	0.78%
E10b-2006	:		:	567	0.78%
SA7-2006	:		:	567	0.78%
SC4-2006	:		:	567	0.78%

Fig. 1 (continued).

CH503034) have been developed and are currently being evaluated in clinical trials (BILN2061 despite reducing viremia 10–1000 fold in patients during phase Ib clinical trial, proved cardiotoxic and its development was halted). However, as with

HIV protease inhibitors, resistance to these inhibitors was expected to become a major drawback of these new antiviral agents. Indeed, treatment with the new inhibitors may induce the appearance of resistant mutants which subsequently would

gain a selective growth advantage over the existing wild-type viruses and rapidly become dominant in the viral quasispecies. This has been demonstrated both *in vitro* (Neyts, 2006) and *in vivo* during phase II clinical trials (Sarrazin et al., 2007). Selective pressure in the replicon system with BILN2061 resulted in several drug-resistant single mutations: A156T, R155Q and D168V (Lu et al., 2004); similar results have been reported with VX-950 for which substitutions A156V/T/S, V36A/M, T54A and R155K/T induced different degrees of resistance (Wohnsland et al., 2007), and also with SCH503034 for which mutations T54A, V170A and A156S conferred low to moderate levels of resistance to SCH503034, while A156T in the replicon system resulted in high-level resistance to this compound and also to the related ketoamide SCH446211 (SCH6) (Yi et al., 2006; Tong et al., 2006a,b).

Here we report that sequences carrying the NS3-protease-inhibitor resistant mutation A156T, were already present (at a frequency of 0.78%) within the liver of an HCV-infected patient who had never been treated with any of these protease inhibitors.

Results

A total of 128 independent clones from the NS3-protease domain (567 nucleotides each) were obtained from the liver biopsy of the anti-protease untreated patient. The population of sequences at the nucleotide level was composed by a master sequence that represented 70.31% (90/128), surrounded by a cloud of mutants, one of them represented 3.91% (5/128), another 2.34% (3/128), another 1.6% (2/128) and a group of single mutants 21.9% (28/128) representing each one 0.78% (1/128) (Fig. 1). The total number of individual mutations was 51 from a total of 72,576 nucleotides sequenced. The master and consensus were identical.

At the amino acid level NS3-protease region was confirmed to be a very well conserved region with 86% (110/128) of sequences equal, representing a master sequence that coincided with the consensus one. The master sequence was surrounded by three sequences with the same mutation that represented 2.3% (3/128) of the population and 11.7% (15/128) of single point mutants representing 0.78% each (Fig. 2). The total number of mutations was 20 from a total of 24,192 amino acids studied.

The principal observation was the presence of mutation A156T that exerts highly effective resistance in replicon system studies to protease inhibitors BILN2061, VX950, SCH446211 and SCH503034 (Yi et al., 2006). In the mutant named B6b-2006 the codon GCC (Ala) changed to ACC (Thr). Mutation in this codon did not seem to be a rare event since another clone B3b-2006 already had a single point mutation but in the third position GCA (Ala) that does not generate an amino acid change (Figs. 1 and 2).

Discussion

Appearance of mutants resistant to anti-NS3-protease treatment has been reported *in vitro* using the replicon system and during treatment, but the presence of such mutants in a patient has not been shown before starting treatment. Here we report

that NS3-protease-inhibitor resistant mutant, A156T, was already present at 0.78% in the infected liver of an HCV-infected patient who had never been treated with any of these protease inhibitors. Previous estimations of HCV production showed an average of 1.3×10^{12} HCV virions produced per day (Neumann et al., 1998), with 4 to 25% (average of 14%) of the 2×10^{11} hepatocytes per liver had replicative-intermediate RNA (Chang et al., 2000), and an average number of viral genomes per productively infected hepatocyte of 33 (7 to 64) (Chang et al., 2003). The total HCV genomic RNA in the liver has been estimated to range from 1.3×10^8 to 1.3×10^9 RNA molecules per g of liver tissue. Following these estimations together with our data that 0.78% of sequences are carrying the protease-inhibitor resistant mutation A156T in our untreated patient, it suggests that at the time of analysis, between 1×10^6 – 1×10^7 RNA molecules per gram of liver tissue may be carrying this particular mutation, which is not negligible.

Mutation frequency is defined as the proportion of mutation positions relative to the consensus nucleotide/amino acid sequence and it is calculated by dividing the number of mutations relative to the consensus by the total number of nucleotide/amino acid sequences. We compared the average mutation frequency of the HCV NS3-protease region with mutation frequency in previously described codons directly or indirectly involved in drug resistance (amino acid V36, T54, R109, R155, A156, D168 and V170 and compensatory mutations Q86, P89 and G162). No statistically significant differences were found. Average mutation frequency and mutation frequencies found in residues of unknown drug resistance was 0.7×10^{-3} at the nucleotide level and 0.83×10^{-3} at the amino acid level, while mutation frequencies at residues related to drug resistance were 0.52×10^{-3} and 0.78×10^{-3} respectively. Therefore, since mutation frequencies at the codons related to resistance are comparable to the average, it is expected that, for statistical reasons, mutations will occasionally affect the drug-resistant codons and relevant substitutions, as the A156T mutation here reported, will appear by random mutation in the absence of selection by the inhibitor. At present, we are not able to differentiate whether naturally occurring A156T mutant was replicating in the liver or was just a simple probabilistic recent event during viral replication of the wild type.

The fact that sequence with mutation A156T represents as much as 0.78% of the liver viral population despite the significant reduction of HCV-NS3/4A catalytic efficiency, poly-protein processing, and replication fitness (Tong et al., 2006a) requires further explanation. Fitness of a viral population depends on multiple factors that are affecting viral replication (from hepatocyte metabolism, to the immune pressure) and this may be different from patient to patient and *in vitro* depending on the cell culture conditions. It has been described that mutations that largely compensate for fitness reduction often arise (Escarmis et al., 1999). It has been recently reported that three separate second-site mutations P89L, Q86R, and G162R were capable of partially reversing A156T defects in polyprotein processing and/or replicon fitness without reducing resistance to protease inhibitor (Yi et al., 2006). Our mutant A156T and also all sequences isolated from liver of patient studied, had changes at positions 89 and 86 (P89Q and Q86P). Since these two

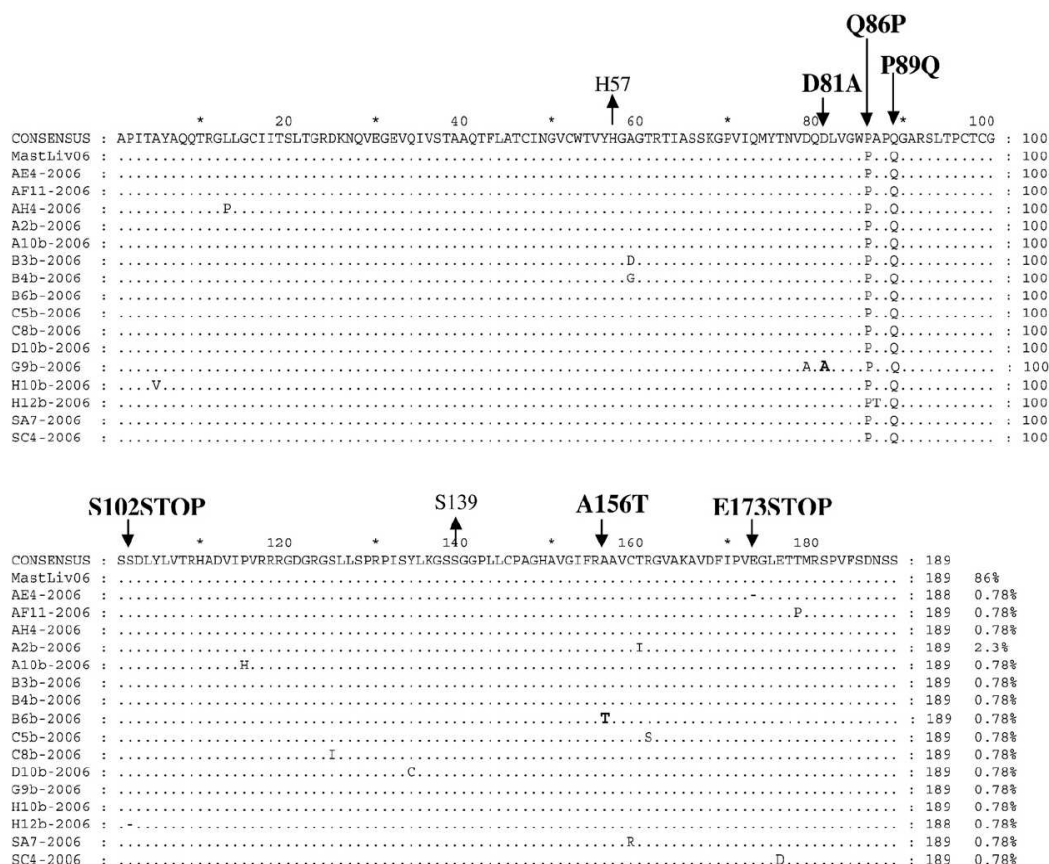


Fig. 2. Alignment of the sequences that were different at the amino acid level. Dots indicate residues identical to those of the consensus sequence. The percentage at which each sequence was present in the isolate is shown at the end of each sequence. Dash indicates STOP mutation. Arrow caps indicate important amino acid changes such as STOP codons, mutation A156T, mutations P89Q, Q86P and mutation D81A that belongs to the catalytic triad H57, D81 and S139. Clonal sequences are identified with the clonal descriptor and the year 2006 when liver biopsy of the chronic patient was obtained.

mutations have been fixed in the quasispecies of our patient, a selective advantage for the mutant carrying them may exist. Taking all together, the most plausible explanation for the presence in 0.78% of the viral population of A156T mutant is that, mutations P89Q and Q86P might compensate for A156T-associated fitness decrease. Therefore, the presence of putative compensatory mutations in all viral liver sequences may have facilitated the random appearance of the A156T mutant, which might otherwise have been negatively selected in the absence of protease-inhibitors treatment. The compensatory fitness effect of these mutations however, would require in vitro confirmation. On the other hand, originally described compensatory mutations P89L and Q86R along with the A156T, if present, would require sequencing 6.4×10^{13} clones to be demonstrated, in the absence of drug treatment.

The presence of drug-resistant RNA viruses from natural isolates in untreated patients is not a rare event and has been previously reported for several other viruses, specially in the case of antiretroviral-resistant HIV-1 (Najera et al., 1995; Mohri et al., 1993), and including acyclovir-resistant herpes simplex virus (Parris and Harrington, 1982), drug-resistant Citomegalovirus (Liu et al., 2000) and also lamivudine-resistant HBV (Kobayashi et al., 2001).

Preexistence of such drug-resistant viruses in an untreated patient is a direct consequence of the quasispecies structure of RNA viruses with the continuous generation of variants, and represents an important drawback in the treatment of viral diseases (Duarte et al., 1994; Domingo and Holland, 1992). In our study, mutation A156T confers high-level resistance to drug inhibitors BILN2061, VX-950, SCH446211 and SCH503034. Therefore any combination therapy using only protease inhibitors, is predicted to be inescapably deemed to failure, opening important ethical considerations in monotherapy treatments and suggest that compounds targeting different regions should be used in combination (Domingo, 1989), to decrease the likelihood of resistance, or in combination with other antivirals like interferon and ribavirin.

Patients and methods

Patient A was a 28-year old man with a history of Intravenous drug use who was found to be anti-HCV positive at the time of a volunteer blood donation in 1993, and prospectively

followed since that time (Quer et al., 2003). He had persistent HCV infection with a genotype 1a as confirmed by sequencing of 5'UTR, E2PePHD and NS3 regions and a viral load in serum of 10^6 IU/mL. A liver biopsy showed moderate chronic active hepatitis and he received treatment with standard interferon and ribavirin (IFN+RBV) for 12 months (between February 1996 and March 1997) with transient biochemical but no virological response. Subsequently, he had not been given further treatment. He underwent a follow up liver biopsy in November 2006 and excess tissue not needed for routine histological assessment was immediately embedded in RNA later (Ambion Cambridge-shire, UK) for 24 h prior to cryopreservation at -80°C .

RNA-extraction, nested RT-PCR using pairs of primers A5319ds 5'CAGGACGCCGCCAACGAGCACCCA3' A3349us 5'TGACATCATCAACGGCTTGCCTGTTTC 3', internal primers and A47us (up): 5'GCGCCCATCACGGCG-TACGCCCA3' and p871ds (down): 5'GCATCCGTG-GAGTGGCACTC3', for nested PCR.

To minimize misincorporation errors during PCR amplification, we used the high-fidelity DNA polymerase Pfu (Stratagene, San Diego, CA, USA) which, under the conditions used, has a reported error rate between 1.3×10^{-6} (Cline et al., 1996) and 6.5×10^{-7} (Andre et al., 1997). After 70 cycles of amplification, the expected number of mutation in the worse reported case is 1 error in 9.1×10^5 nucleotides sequenced. In our study, after sequencing 45,432 nucleotides, the expected number of errors due to Pfu is 0.5. Besides, a control study has been carried out and one of the clones (1893 nucleotides) of known sequence, was subjected to 70 additional cycles of PCR in 20 independent amplifications. No mutations were detected after analyzing 8140 nucleotides. This result improves our previous estimates of the misincorporation rate when using a more error-prone polymerase, Taq DNA polymerase (Perkin-Elmer) (1/5451 and 2/9964) (Cabot et al., 2001; Martell et al., 1992), and confirms that most of the observed heterogeneity was independent of artifact during the amplification procedure (Quer et al., 2005). GenBank accession numbers for different sequences are EF613694 to EF613725.

Acknowledgments

This work has been supported in part by the Spanish Ministry of Science and Education grant SAF 2006-03681; the Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III, grant PI061244, Fundació Marató TV3 052310 and FIPSE 36623/06. M. Cubero has a predoctoral fellowship from Fundació Marató TV3 052310.

References

- Andre, P., Kim, A., Khrapko, K., Thilly, W.G., 1997. Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. *Genome Res.* 7, 843–852.
- Cabot, B., Martell, M., Esteban, J.I., Piron, M., Otero, T., Esteban, R., Guardia, J., Gomez, J., 2001. Longitudinal evaluation of the structure of replicating and circulating hepatitis C virus quasispecies in nonprogressive chronic hepatitis C patients. *J. Virol.* 75, 12005–12013.
- Chang, M., Marquardt, A.P., Wood, B.L., Williams, O., Cotler, S.J., Taylor, S.L., Carithers Jr., R.L., Gretch, D.R., 2000. In situ distribution of hepatitis C virus replicative-intermediate RNA in hepatic tissue and its correlation with liver disease. *J. Virol.* 74, 944–955.
- Chang, M., Williams, O., Mittler, J., Quintanilla, A., Carithers Jr., R.L., Perkins, J., Corey, L., Gretch, D.R., 2003. Dynamics of hepatitis C virus replication in human liver. *Am. J. Pathol.* 163, 433–444.
- Cline, J., Braman, J.C., Hogrefe, H.H., 1996. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* 24, 3546–3551.
- Domingo, E. (1989). RNA virus evolution and the control of viral disease. In "Progress in drug research." (Ed. Jucker Birkhauser Verlag, Ed.), Vol. 33, pp. 93–133. Vol. XXXIII. Basel.
- Domingo, E., Holland, J.J., 1988. High error rates, population equilibrium, and evolution of RNA replication systems. In: Domingo, E., Holland, J.J. (Eds.), *RNA Genetics*. CRC Press, Boca Raton, Florida, pp. 3–36.
- Domingo, E., Holland, J.J., 1992. Complications of RNA heterogeneity for the engineering of virus vaccines and antiviral agents. *Genet. Eng. (N.Y.)* 14, 13–31.
- Domingo, E., Holland, J.J., 1994. Mutation rates and rapid evolution of RNA viruses. In: Morse, S.S. (Ed.), *Evolutionary Biology of Viruses*. Raven Press, New York, pp. 161–184.
- Duarte, E.A., Novella, I.S., Weaver, S.C., Domingo, E., Wain-Hobson, S., Clarke, D.K., Moya, A., Elena, S.F., De La Torre, J.C., Holland, J.J., 1994. RNA virus quasispecies: significance for viral disease and epidemiology. *Infect. Agents Dis.* 3, 201–214.
- Eigen, M., Biebricher, C.K., 1988. Variability of RNA genomes. Sequence space and quasispecies distribution. In: Domingo, E., Holland, J.J., Ahlquist, P. (Eds.), *RNA Genetics*, vol. 3. CRC Press, Inc., Boca Raton, Florida, pp. 211–245.
- Escarmis, C., Davila, M., Domingo, E., 1999. Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J. Mol. Biol.* 285, 495–505.
- Fried, M.W., Shiffman, M.L., Reddy, K.R., Smith, C., Marinos, G., Goncalves Jr., F.L., Haussinger, D., Diago, M., Carosi, G., Dhumeaux, D., Craxi, A., Lin, A., Hoffman, J., Yu, J., 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.* 347, 975–982.
- Holland, J.J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S., 1982. Rapid evolution of RNA genomes. *Science* 215, 1577–1585.
- Holland, J.J., De La Torre, J.C., Steinhauer, D.A., 1992. RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* 176, 1–20.
- Kobayashi, S., Ide, T., Sata, M., 2001. Detection of YMDD motif mutations in some lamivudine-untreated asymptomatic hepatitis B virus carriers. *J. Hepatol.* 34, 584–586.
- Kolykhalov, A.A., Mihalik, K., Feinstone, S.M., Rice, C.M., 2000. Hepatitis C virus-encoded enzymatic activities and conserved ARN elements in the 3' nontranslated-region are essential for virus replication in vivo. *J. Virol.* 74, 2046–2051.
- Liu, W., Shum, C., Martin, D.F., Kuppermann, B.D., Hall, A.J., Margolis, T.P., 2000. Prevalence of antiviral drug resistance in untreated patients with cytomegalovirus retinitis. *J. Infect. Dis.* 182, 1234–1238.
- Lu, L., Pilot-Matias, T.J., Stewart, K.D., Randolph, J.T., Pithawalla, R., He, W., Huang, P.P., Klein, L.L., Mo, H., Molla, A., 2004. Mutations conferring resistance to a potent hepatitis C virus serine protease inhibitor in vitro. *Antimicrob. Agents Chemother.* 48, 2260–2266.
- Manns, M.P., McHutchison, J.G., Gordon, S.C., Rustgi, V.K., Shiffman, M.L., Reindollar, R., Goodman, Z.D., Koury, K., Ling, M.H., Albrecht, J.K., 2001. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* 358, 958–965.
- Martell, M., Esteban, J.I., Quer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J., Gomez, J., 1992. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J. Virol.* 66, 3225–3229.
- Mohri, H., Singh, M.K., Ching, W.T., Ho, D.D., 1993. Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the blood of treated and untreated patients. *Proc. Natl. Acad. Sci. U. S. A.* 90, 25–29.
- Najera, I., Holguin, A., Quinones-Mateu, M.E., Munoz-Fernandez, M.A., Najera, R., Lopez-Galindez, C., Domingo, E., 1995. Pol gene quasispecies of

- human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy. *J. Virol.* 69, 23–31.
- Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J., Perelson, A.S., 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* 282, 103–107.
- Neyts, J., 2006. Selective inhibitors of hepatitis C virus replication. *Antivir. Res.* 71, 363–371.
- Parris, D.S., Harrington, J.E., 1982. Herpes simplex virus variants restraint to high concentrations of acyclovir exist in clinical isolates. *Antimicrob. Agents Chemother.* 22, 71–77.
- Quer, J., Esteban, J.I., 2005. Epidemiology. In: Thomas, H.C., Lemon, S., Zuckerman, A.J. (Eds.), *Viral Hepatitis*. Blackwell Publishing, Oxford, pp. 407–425.
- Quer, J., Murillo, P., Esteban, J.I., Martell, M., Esteban, R., Guardia, J., 2003. Sexual transmission of hepatitis C virus from a patient with chronic disease to his sex partner after removal of an intrauterine device. *Sex. Transm. Dis.* 30, 470–471.
- Quer, J., Esteban, J.I., Cos, J., Sauleda, S., Ocana, L., Martell, M., Otero, T., Cubero, M., Palou, E., Murillo, P., Esteban, R., Guardia, J., 2005. Effect of bottlenecking on evolution of the nonstructural protein 3 gene of hepatitis C virus during sexually transmitted acute resolving infection. *J. Virol.* 79, 15131–15141.
- Sarrazin, C., Kieffer, T.L., Bartels, D., Hanzelka, B., Muh, U., Welker, M., Wincheringer, D., Zhou, Y., Chu, H.M., Lin, C., Weegink, C., Reesink, H., Zeuzem, S., Kwong, A.D., 2007. Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology* 132, 1767–1777.
- Steinhauer, D.A., Holland, J.J., 1987. Rapid evolution of RNA viruses. *Annu. Rev. Microbiol.* 41, 409–433.
- Tong, X., Chase, R., Skelton, A., Chen, T., Wright-Minogue, J., Malcolm, B.A., 2006a. Identification and analysis of fitness of resistance mutations against the HCV protease inhibitor SCH 503034. *Antivir. Res.* 70, 28–38.
- Tong, X., Guo, Z., Wright-Minogue, J., Xia, E., Prongay, A., Madison, V., Qiu, P., Venkatraman, S., Velazquez, F., Njoroge, F.G., Malcolm, B.A., 2006b. Impact of naturally occurring variants of HCV protease on the binding of different classes of protease inhibitors. *Biochemistry* 45, 1353–1361.
- Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R., 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439, 344–348.
- Wohnsland, A., Hofmann, W.P., Sarrazin, C., 2007. Viral determinants of resistance to treatment in patients with hepatitis C. *Clin. Microbiol. Rev.* 20, 23–38.
- Yi, M., Tong, X., Skelton, A., Chase, R., Chen, T., Prongay, A., Bogen, S.L., Saksena, A.K., Njoroge, F.G., Veselenak, R.L., Pyles, R.B., Bourne, N., Malcolm, B.A., Lemon, S.M., 2006. Mutations conferring resistance to SCH6, a novel hepatitis C virus NS3/4A protease inhibitor. Reduced RNA replication fitness and partial rescue by second-site mutations. *J. Biol. Chem.* 281, 8205–8215.

ESTUDIO 3

5 ESTUDIO 3: IDENTIFICACIÓN DE FACTORES VIRALES Y DEL PACIENTE INVOLUCRADOS EN RESOLUCIÓN O CRONICIDAD

5.1 INTRODUCCIÓN

Después de la infección por el VHC, el curso natural de la infección hacia la resolución espontánea o cronicidad depende de:

5.1.1 Factores virales (huésped)

5.1.2 Factores genómicos del paciente (hospedador).

5.2 OBJETIVOS

El objetivo de esta última parte fue el estudio de factores virales y del paciente que podrían ayudar a explicar la evolución hacia cronicidad o resolución espontánea, mediante secuenciación masiva con la plataforma 454-FLX (Roche Diagnostics). Debido a la importancia del tamaño del inóculo viral que se transmite, se estudiaron 3 rutas diferentes de transmisión del VHC comparando la fuente (paciente crónico) y el receptor: **(i)** transmisión de un número bajo de partículas por cuello de botella genético (pacientes A y B), **(ii)** transmisión de un número indeterminado de virus (pacientes C y D) **(iii)** la transmisión masiva de partículas desde un paciente antes y después del trasplante de hígado (paciente Epre Y Epost). La descripción de pacientes se encuentra en el apartado 8.1 de Materiales y Métodos. Se estudiaron 6 aislados de suero de 5 pacientes.

5.3 RESULTADOS

Factores genómicos del paciente

5.3.1 Haplotipos HLA

El paciente B presentó 3 haplotipos MHC clase II y 1 haplotipo MHC clase I asociados con resolución de la infección (Tabla 3). Los alelos de clase II DRB1*1101, DQB1*0301 y DQB1*0302 se han relacionado con resolución espontánea de la infección (96;98;101). El paciente A presentó 2 haplotipos de HLA clase II sobre los cuales existe cierta discrepancia sobre su implicación en curación, el HLA DQB1*0501 que se asoció con resolución de la infección en una cohorte de pacientes de Irlanda (112) y el haplotipo HLA-DRB1*1501, sobre el cual se indentificó un epítipo Th1 (1388-1407) restringido a este HLA, donde mutantes de escape en esta región, pueden atenuar o inhibir la proliferación de las células T (89;102). El paciente D presentó 2 haplotipos de clase II que se han relacionado con persistencia de la infección (101). En este estudio no se disponen de datos de haplotipos de HLA de los paciente C y E.

Tabla 3. Haplotipos de HLA de los pacientes A, B, C, D y E.

Haplotipos de HLA					
Pacientes	HLA de clase I			HLA de clase II	
	HLA A	HLA B	HLA C	HLA-DR	HLA-DQ
Paciente A (crónico)	A*201	B*0702	Cw*0702	DRB1*0102 DRB1*1501 HLA-DRB5*	DQB1*0501 DQB1*0602
Paciente B (curado)	A*6601 A*6901	B*4102 B*5501	Cw*0102 Cw*1701	DRB1*0402 DRB1*1101 DRB3*02 DRB4*	DQB1*0301 DQB1*0302
Paciente C (crónico)	ND				
Paciente D (inicio de cronicidad)	A*02 A*25	B*18		DRB1*03 DRB1*13	DQB1*02 DQB1*06
Paciente E (paciente transplantado)	ND				

5.3.2 Genotipo IL28B

Los pacientes agudos, B y D fueron homocigotos para el genotipo favorable de IL28 CC en el SNP rs12979860, mientras que el paciente crónico A fue heterocigoto (CT) y el paciente E presentó un genotipo TT desfavorable (78;79). Importante destacar que el paciente D, a pesar de ser CC, el virus había conseguido superar a la respuesta inmune y estaba en vías de cronificar.

Tabla 4. SNP “upstream” del gen IL28B

SNP IL28B	GENOTIPOS ASOCIADOS A CURACIÓN ESPONTÁNEA O SVR	PACIENTE A (crónico)	PACIENTE B (curación)	PACIENTE C (crónico)	PACIENTE D (crónico)	PACIENTE E (crónico)
rs12979860	CC	CT	CC	ND	CC	TT

En resumen, el paciente B curado, tenía el genotipo favorable para el polimorfismo rs12979860 del gen IL28 y haplotipos de HLA clase I y II asociados con curación. El paciente crónico A tenía un genotipo intermedio CT de IL28 y algún haplotipo favorable de HLA. El paciente D también poseía el genotipo CC favorable de IL28, pero presentó polimorfismos de HLA relacionados con persistencia. El paciente E presentó el genotipo desfavorable TT en IL28 (HLA desconocido). Estos datos serían indicativos de que el paciente B tiene más probabilidades de curación, que los pacientes A, D y E. Los datos genéticos, tanto de IL28 como HLA, deben traducirse en una respuesta funcional (fenotipo), es decir, en un efecto sobre la respuesta inmune y sobre la dinámica de la población viral.

Factores virales

La complejidad viral en los pacientes incluidos en este estudio se caracterizó por una serie de parámetros (descrito en materiales y métodos apartado 8.3.6 página 139) que permitieron el desarrollo de estudios comparativos entre poblaciones. Se analizó el gen NS3 para cubrir toda la región (y teniendo en cuenta que el GS-FLX química Titanium permite la secuenciación de fragmentos de 400 nucleótidos) se dividió la NS3 en 7 amplicones superpuestos (ver en pacientes materiales y métodos Figura 22) obtenidos utilizando cebadores específicos (ver en pacientes Materiales y Métodos Figura 22 Tabla 16). Un amplicón es un fragmento de ADN o ARN que se amplifica mediante unos cebadores específicos. Las amplificaciones se enumeraron del 1 al 7 desde el extremo 5' al extremo 3' de la región codificante NS3.

5.3.3 Análisis a nivel de nucleótidos de la complejidad viral de la quasiespecie

La complejidad de las poblaciones virales se representó por los parámetros de diversidad nucleotídica (P_i), frecuencia de mutación (M_f) y entropía de Shannon normalizada (S_n) para los 6 pacientes incluidos en este estudio (Tabla 5).

El análisis de complejidad de la población viral de la pareja de paciente A y B a nivel de nucleótidos mostró como, después de la transmisión por cuello de botella y en la fase aguda, se produjo una homogenización importante de la población viral con una diferencia de una orden de magnitud (por ejemplo, los valores promedios de P_i fueron $5,9E-03$ v $6,13E-04$; $M_f = 4,125E-03$ v $3,33E-04$). La complejidad no se distribuyó de la misma manera lo largo de la región NS3 (ver amplicón por amplicón en Tabla 5).

En el caso del paciente E, después del trasplante, la disminución de la complejidad fue de $1/3$, mientras que en los pacientes C y D, tan sólo disminuyó a la mitad.

Por tanto, en las 3 situaciones clínicas diferentes, tan sólo el paciente B que se curó de manera espontánea, presentó una reducción importante de la complejidad (1 log).

La complejidad no fue homogénea en los 7 amplicones de cada uno de los pacientes estudiados (Tabla 5). El estudio de complejidad abarcando fragmentos nucleotídicos grandes puede enmascarar que la variabilidad se concentre en subregiones del genoma estudiado. Para determinar con mayor exactitud cómo se distribuía la variabilidad a lo largo de un segmento genómico, se realizó el estudio de complejidad por ventanas de 50 nts que van corriendo de 10 en 10 nts a lo largo del genoma (en cada paso se solapan 40 nts con la ventana anterior).

Tabla 5. Indicadores de complejidad de quasiespecies (diversidad de nucleótidos, frecuencia de mutación y entropía normalizada de Shannon) por amplicón a nivel de nucleótidos. * En el amplicón 1 del paciente D no existe un número suficiente de secuencias para realizar el estudio de complejidad de la población viral.

Diversidad nucleotídica (Pi)									
Amplicones	A	B	A/B	C	D	C/D	Epre	Epost	Epre/Epost
3297-3634	1,005E-02	2,284E-04	44,00	2,995E-03	0,000E+00	*	9,358E-03	4,011E-03	2,33
3555-3944	5,542E-03	5,034E-04	11,01	2,952E-03	2,017E-03	1,46	1,265E-02	3,224E-03	3,92
3840-4196	8,790E-03	5,097E-04	17,25	4,132E-03	1,526E-03	2,71	3,251E-02	5,936E-03	5,48
4062-4391	1,308E-02	9,867E-04	13,26	4,500E-03	1,886E-03	2,39	2,302E-02	6,147E-03	3,74
4302-4661	4,613E-04	6,366E-04	0,72	3,613E-03	2,175E-03	1,66	1,295E-02	4,520E-03	2,87
4614-4986	7,219E-04	1,002E-03	0,72	1,736E-03	8,309E-04	2,09	1,247E-02	6,279E-03	1,99
4926-5281	2,761E-03	4,282E-04	6,45	2,483E-03	6,825E-04	3,64	1,009E-02	7,148E-03	1,41
media	5,915E-03	6,136E-04	9,64	3,236E-03	1,520E-03	2,13	1,615E-02	5,324E-03	3,03

Frecuencia de sustitución de nucleótidos (Mf)									
Amplicones	A	B	A/B	C	D	C/D	Epre	Epost	Epre/Epost
3297-3634	7,227E-03	1,153E-04	62,68	1,542E-03	0,000E+00	*	5,195E-03	2,190E-03	2,3720
3555-3944	3,970E-03	2,539E-04	15,64	1,526E-03	1,061E-03	1,44	7,039E-03	1,816E-03	3,8755
3840-4196	5,664E-03	2,574E-04	22,00	2,430E-03	8,013E-04	3,03	2,316E-02	3,492E-03	6,6327
4062-4391	9,843E-03	5,039E-04	19,53	2,657E-03	9,760E-04	2,72	1,505E-02	4,428E-03	3,3999
4302-4661	2,333E-04	3,370E-04	0,69	2,233E-03	1,400E-03	1,59	7,832E-03	3,119E-03	2,5111
4614-4986	3,641E-04	6,343E-04	0,57	8,846E-04	4,202E-04	2,11	8,153E-03	4,126E-03	1,9762
4926-5281	1,576E-03	2,312E-04	6,82	1,271E-03	3,471E-04	3,66	6,382E-03	4,691E-03	1,3604
media	4,125E-03	3,333E-04	12,38	1,834E-03	8,343E-04	2,20	1,040E-02	3,409E-03	3,0515

Entropía de Shannon Normalizada (Sn)									
Amplicones	A	B	A/B	C	D	C/D	Epre	Epost	Epre/Epost
3297-3634	0,7230	0,1101	6,57	0,323371	0,000000	*	0,591691	0,497066	1,19
3555-3944	0,6095	0,2079	2,93	0,424087	0,371677	1,14	0,654525	0,508232	1,29
3840-4196	0,5864	0,1813	3,23	0,440493	0,225281	1,96	0,728917	0,622167	1,17
4062-4391	0,6553	0,3029	2,16	0,395127	0,175743	2,25	0,744650	0,521714	1,43
4302-4661	0,1616	0,2147	0,75	0,430541	0,326234	1,32	0,723422	0,478055	1,51
4614-4986	0,2563	0,3774	0,68	0,310540	0,146253	2,12	0,703957	0,567470	1,24
4926-5281	0,4770	0,2745	1,74	0,383644	0,182148	2,11	0,710613	0,636687	1,12
media	4,956E-01	2,384E-01	2,08	0,397405	0,152352	2,61	0,693968	0,547342	1,27

Sliding Window de nucleótidos

El análisis del gen NS3 del VHC se realizó mediante un “sliding window” (detallado en el apartado de materiales y métodos apartado 8.3.6. página 147), usando un tamaño de ventana de 50, y un tamaño de pasos de 10 nucleótidos. El resultado global en todo NS3 se representa en la Figura 5.

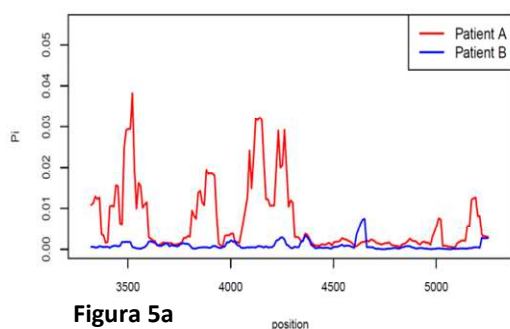


Figura 5a

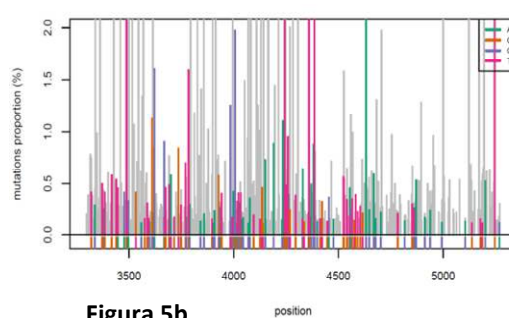


Figura 5b

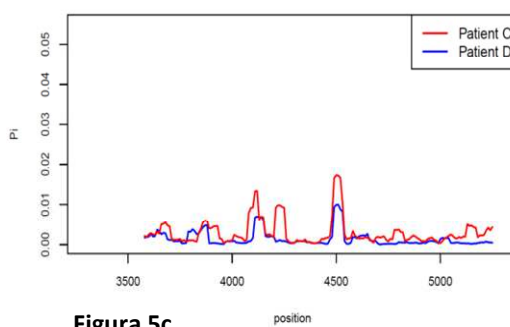


Figura 5c

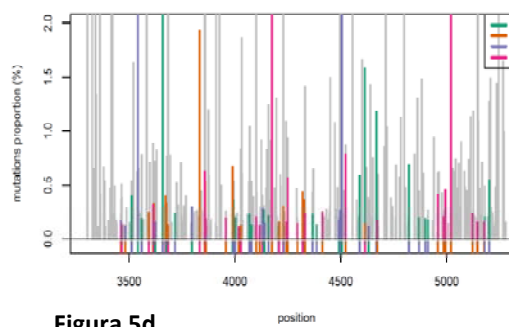


Figura 5d

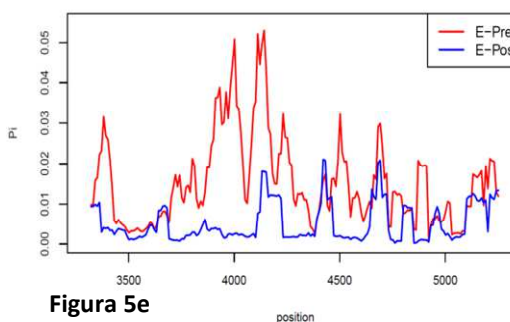


Figura 5e

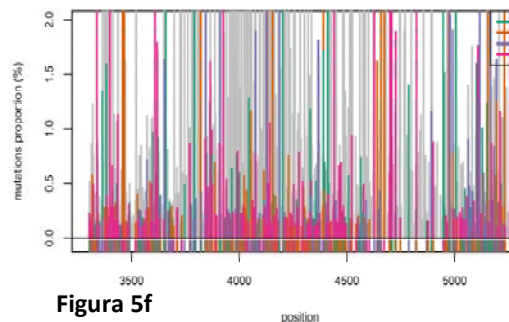


Figura 5f

Figura 5. a, c y e) Comparación de la diversidad de nucleótidos de los pacientes A-B, C-D y Epre-Epost, por sliding windows de 50nt en pasos de 10nt, donde en abscisas (y) están representados los valores de diversidad nucleotídica (P_i) y en el eje de ordenadas (x) las posiciones de nucleótidos a lo largo de la proteína NS3 del VHC. La diversidad nucleotídica tiene en cuenta el número medio de mutaciones entre cada par de secuencias. **b, d y f)** Población de sustituciones de nucleótidos posición por posición, donde se representó en el eje de abscisas (y) la proporción de mutaciones con respecto a la posición nucleotídica en el eje de ordenadas. Para distinguir las situaciones observadas en cada pareja estudiada, los aislados de los pacientes B, D y Epost están representados en colores diferentes y los pacientes A, C y Epre en gris.

El paciente crónico A presentó una variabilidad nucleotídica mayor en 3 regiones de la proteasa y en el extremo 3' de la helicasa, con respecto al paciente agudo B (Figura 5a). Este paciente B mostró una baja diversidad a lo largo de toda la secuencia de NS3. En el caso de los pacientes C (crónico) y D (agudo) la diversidad nucleotídica fue baja y similar entre ambos, sin grandes diferencias (excepto nucleótidos puntuales) tanto en la proteasa como en la helicasa (Figura 5c). En la tercera situación, el paciente E antes del trasplante de hígado, presentó una gran diversidad nucleotídica que se distribuyó a lo largo del gen NS3. Después del trasplante hepático, la variabilidad fue menor, pero con picos altos de variación tanto en la proteasa como en la helicasa (Figura 5e).

La distribución de la variabilidad del VHC en NS3 posición por posición (Figuras 5b, d y f), mostró que en los pacientes crónicos (representados en gris) existió un mayor porcentaje de mutaciones, que en los pacientes agudos. De entre los 3 aislados de pacientes crónicos, el Epre fue el que presentó una variabilidad mayor.

5.3.4 Análisis a nivel de aminoácidos de la complejidad viral de la quasiespecie

En la Tabla 6 se representan los valores de diversidad, frecuencia de mutación y entropía de Shannon a nivel de aminoácidos. En el caso de los pacientes A y B, en contraste a las secuencias de nucleótidos, la complejidad a nivel de aminoácidos fue diferente, mientras que el paciente A tenía una diversidad de nucleótidos mayor que B en los 7 amplicones estudiados, ambos pacientes presentaron una diversidad, S_n y Frecuencia de mutación similar a nivel de aminoácidos. Por ejemplo, la diversidad a nivel de aminoácidos para el paciente A fue de $0,819E-03$ similar a la del B= $1,09E-03$. El cociente A/B de diversidad pasó de 9,64 en nucleótidos a 0,75 en aminoácidos (Tabla 6).

En general, los valores de P_i a nivel de aminoácidos en cada amplicón de los pacientes C, D, Epre y Epost, prácticamente no varió del observado a nivel de nucleótidos (de 2,13 a 1,22 y de 3,03 a 2,439). Se observaron resultados similares para la M_f y S_n .

En los pacientes A y B, el amplicón más interesante fue el 7, puesto que se comportó diferente al resto de amplicones. En el paciente A, los amplicones de NS3 del 2 al 6 (el 1 incluye un fragmento de NS2 y se descarta de la discusión) mostraron una disminución significativa en la diversidad respecto al observado a nivel de nucleótidos, indicando que la mayoría de cambios fueron sinónimos (por ejemplo, el amplicón 3 con un valor de $8.79\text{E-}03$ baja a $9.19\text{E-}05$). Por el contrario, el amplicón 7 en el paciente A prácticamente no varía y pasa de $2,7\text{E-}03$ a $1,15\text{E-}03$, indicando que una gran parte de las mutaciones fueron no sinónimas. En el paciente B, a diferencia de A, prácticamente no varió la diversidad en los amplicones del 2 al 6 (por ejemplo en el amplicón 3 el valor de diversidad nucleotídica fue de $5,09\text{E-}04$ y a nivel de aminoácidos fue de $9,8\text{E-}04$). En el paciente B, el amplicón 7, al contrario del resto, presentó una gran homogeneización bajando a niveles de variación de aa indetectables. Todos estos resultados indican un comportamiento especial en el amplicón 7, alertando de una posible implicación de la variación antigénica en esta región carboxi-terminal de NS3 con el resultado de la infección.

Si comparamos los pacientes con nueva infección A, D y Epost, la homogeneización en el amplicón 7, tan sólo se observó en el paciente que resolvió la infección (paciente B).

Tabla 6. Indicadores de complejidad de quasiespecies (diversidad, frecuencia de mutación y entropía normalizada de Shannon) a nivel de aminoácidos. * En el amplicón 1 del paciente D no existe un número suficiente de secuencias para realizar el estudio de complejidad de la población viral. #En el amplicón 7 del paciente B no se observó cambio alguno de aa.

Diversidad aminoácidos (Pi)									
Amplicones	A	B	A/B	C	D	C/D	Epre	Epost	Epre/Epost
986-1097	4,074E-03	1,192E-04	34,18	4,497E-03	4,742000E-04	*	9,151633E-03	3,358394E-03	2,73
1072-1201	0,000E+00	1,039E-03	0,00	4,190E-03	3,228445E-03	1,30	7,309162E-03	1,772841E-03	4,12
1167-1285	9,195E-05	9,888E-04	0,09	3,163E-03	2,482748E-03	1,27	1,724158E-02	8,224166E-03	2,10
1241-1350	7,425E-05	9,422E-04	0,08	2,226E-03	2,106872E-03	1,06	1,418490E-02	7,661772E-03	1,85
1321-1440	5,813E-05	1,544E-03	0,04	3,187E-03	2,144121E-03	1,49	1,568364E-02	1,667821E-03	9,40
1425-1548	2,810E-04	3,015E-03	0,09	1,230E-03	1,771326E-03	0,69	9,753233E-03	3,390303E-03	2,88
1529-1646	1,156E-03	0,000E+00	#	9,304E-04	4,980277E-04	1,87	4,814623E-03	6,087194E-03	0,79
Mean	8,193E-04	1,093E-03	0,75	2,487964E-03	2,038590E-03	1,22	1,116268E-02	4,594642E-03	2,43

Frecuencia de sustitución de aminoácidos (Mf)									
Amplicones	A	B	A/B	C	D	C/D	Epre	Epost	Epre/Epost
986-1097	3,144E-03	5,981E-05	52,57	2,288E-03	2,374000E-04	*	5,205337E-03	1,695847E-03	3,07
1072-1201	0,000E+00	5,247E-04	0,00	2,150E-03	1,708027E-03	1,26	4,089428E-03	8,960052E-04	4,56
1167-1285	4,622E-05	5,008E-04	0,09	1,639E-03	1,299570E-03	1,26	1,225205E-02	5,560623E-03	2,20
1241-1350	3,727E-05	4,790E-04	0,08	1,138E-03	1,093808E-03	1,04	9,408458E-03	5,738173E-03	1,64
1321-1440	2,917E-05	8,276E-04	0,04	1,651E-03	1,092798E-03	1,51	9,365109E-03	8,366104E-04	11,19
1425-1548	1,411E-04	1,908E-03	0,07	6,173E-04	8,984723E-04	0,69	6,564774E-03	1,817816E-03	3,61
1529-1646	5,976E-04	0,000E+00	#	4,666E-04	2,495213E-04	1,87	2,721225E-03	3,737945E-03	0,73
media	5,708E-04	6,143E-04	0,93	1,277001E-03	1,057033E-03	1,21	7,086625E-03	2,897574E-03	2,45

Entropía de Shannon aminoácidos (Sn)									
Amplicones	A	B	A/B	C	D	C/D	Epre	Epost	Epre/Epost
986-1097	0,9360	0,0408	22,97	0,214812	0,032580	*	0,316305	0,220133	1,44
1072-1201	0,0000	0,1689	0,00	0,278718	0,272274	1,02	0,311836	0,161184	1,93
1167-1285	0,0492	0,1465	0,34	0,202814	0,187696	1,08	0,466901	0,348321	1,34
1241-1350	0,0384	0,1036	0,37	0,121501	0,104954	1,16	0,425438	0,279842	1,52
1321-1440	0,0336	0,2204	0,15	0,208157	0,182279	1,14	0,481250	0,146854	3,28
1425-1548	0,0698	0,3774	0,18	0,118700	0,124040	0,96	0,372572	0,201607	1,85
1529-1646	0,1642	0,0000	#	0,095051	0,072145	1,32	0,272814	0,293432	0,93
media	1,845E-01	1,511E-01	1,22	0,170824	0,157231	1,09	0,378159	0,235910	1,60

Sliding window a nivel de aminoácidos

Al igual que para nucleótidos, el comportamiento diferente de los amplicones a nivel de aminoácidos, indujo a realizar el estudio por ventanas discretas (“sliding windows”) donde el tamaño de ventana fue de 12 (posible tamaño de un epítipo) y la amplitud de pasos de 3.

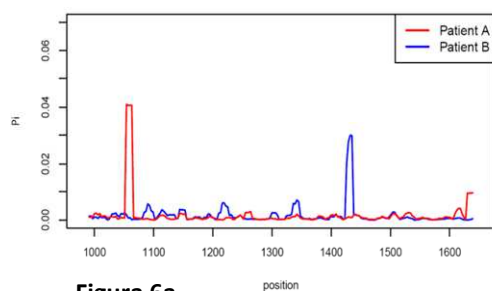


Figura 6a

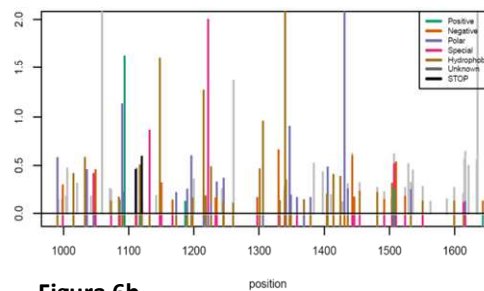


Figura 6b

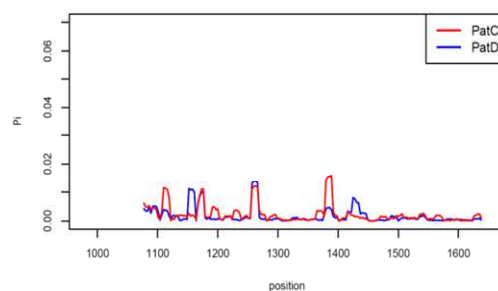


Figura 6c

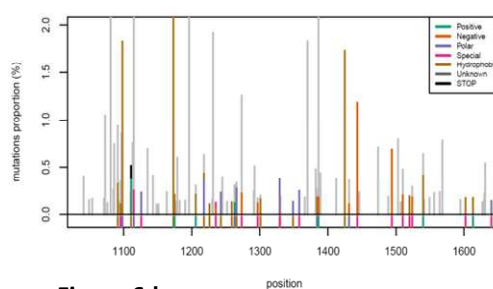


Figura 6d

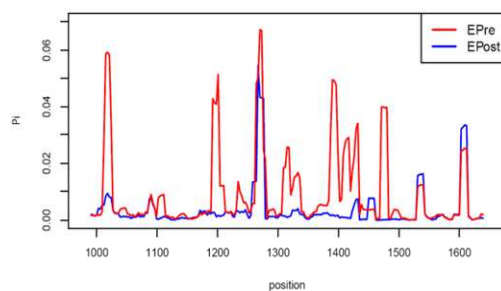


Figura 6e

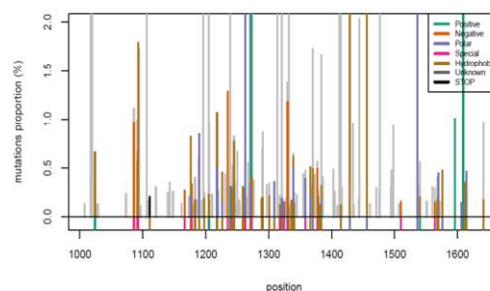


Figura 6f

Figura 6. A, c y e) Comparación de la diversidad de aminoácidos de los pacientes A-B, C-D, y E-pre y E-post, por sliding windows de 12nt en pasos de 3nts. **b, d y f)** Población de sustituciones de aminoácidos posición por posición (los pacientes A, C y Epre están representados en gris y los pacientes B, D y Epost están representados en colores diferentes. En el eje de ordenadas se representaron las posiciones nucleotídicas del gen NS3 y en el eje de abscisas se representó el porcentaje de mutación de cada aminoácido dentro de la población (coloreado por tipo de cambio según propiedades químicas de cada aminoácido).

La diversidad de la población viral a nivel de aminoácidos representada en las gráficas de sliding window (Figura 6) refleja la diferencia en el patrón de cambios respecto al estudio a nivel de nucleótidos (Figura 5). Descartando las 2 posiciones polimórficas (paciente A=Ile1059Val; paciente B=Asp1431Asn) observadas en la gráfica (Figura 6a), la diversidad de aminoácidos fue similar entre A y B, exceptuando el extremo C-terminal de la helicasa, donde los cambios de aminoácidos fueron más prevalentes en el paciente A. En los pacientes C y D, y Epre y Epost, el patrón de cambios de aminoácidos (Figura 6) fue similar al observado a nivel de nucleótidos (Figura 5).

En el estudio de las sustituciones de aminoácidos posición por posición (Figuras 6b, d y f), de los 6 pacientes, la muestra Epre fue la que presentó un mayor porcentaje de mutaciones no sinónimas.

5.3.5 dN/dS poblacional

La relación entre cambios sinónimos y no sinónimos se representó por el cociente de mutaciones no sinónimas respecto a las sinónimas.

De la observación de los valores promedio dN/dS , se concluyó que en el paciente que resolvió la infección (paciente B) predominaron las mutaciones no sinónimas ($dN > dS$) con un valor de sustitución cercano a 1 (0.77), este resultado se interpreta como un predominio de sustituciones por azar. La segunda población más homogénea fue la del paciente D con un valor promedio de dN/dS de 0,47. Este paciente se encontraba en una fase de rebrote viral que, posteriormente, condujo a cronicidad. En el resto de los casos, tanto los pacientes crónicos A, C y Epre, como la situación de Epost-transplante, que como se ha visto anteriormente, presentó un patrón de complejidad parecido a la situación crónica, en todos ellos el valor dN/dS presentó un valor lejos del valor de azar ($dN/dS \ll 1$), por tanto, predominaron las mutaciones sinónimas, y este hecho se ha asociado a una situación de selección natural.

Tabla 7. Relación entre los valores medios de dN (mutaciones no-sinónimas por posiciones no-sinónimas) y dS (mutaciones sinónimas por posiciones sinónimas) que resulta de la comparación de todos los pares de secuencias en la población.

* No determinado debido a escasas diferencias poblacionales en el amplicón.

Amplicones	dN/dS poblacional					
	PACIENTE A	PACIENTE B	PACIENTE C	PACIENTE D	PACIENTE Epre	PACIENTE Epost
3297-3634	0,08344	0,11640	0,48560	0,16459	0,26024	0,27192
3555-3944	0,00000	1,16800	0,43587	0,57274	0,11383	0,12139
3840-4196	0,00150	0,83710	0,16353	0,56891	0,10104	0,44510
4062-4391	0,00077	0,22700	0,10990	0,29642	0,12226	0,37386
4302-4661	0,02354	2,27300	0,21931	0,29621	0,36206	0,07536
4614-4986	0,06878	*	0,15216	1,28822	0,18401	0,10532
4926-5281	0,08356	0,00000	0,06434	0,11586	0,09422	0,19842
Media	0,03737	0,77025	0,23296	0,47185	0,17681	0,22734

5.3.6 Interacción entre factores virales y del paciente. Variabilidad y respuesta inmune adaptativa CD4 específica

El genotipo de HLA está directamente relacionado con la respuesta inmune. En pacientes con resolución espontánea de la infección se ha demostrado una respuesta de células T CD4 múltiple y vigorosa a péptidos de NS3 (113). La respuesta de células T frente a antígenos de VHC restringidos a HLA de clase II se ha asociado con resultado de la infección por VHC, porque se ha descrito que es fuerte en pacientes que resuelven la infección aguda por VHC, mientras que es débil en pacientes que desarrollan una infección crónica (87;95).

En nuestro estudio, se tuvo acceso a la donación de PBMC por parte de los pacientes A y B, lo cual permitió la evaluación de la respuesta CD4 específica mediante ELISPOT, donde se midió el número de células CD4 específicas productoras de IFN- γ frente a pools de 13 péptidos que cubrían el gen NS3 helicasa.

El paciente crónico A presentó una respuesta inmune CD4 débil y concentrada en 2 epítomos (**aa1241-aa1274** y **aa1583-aa1604**) restringidos por el haplotipo HLA DRB1*1501.

El paciente B, que resolvió la infección, presentó una respuesta inmune CD4+ fuerte y amplia frente a epítomos restringidos para su haplotipo HLA DRB1*1101 (**aa1423-aa1458, aa1459-1520**), y también frente a 2 epítomos (**aa1583-aa1604 y aa1634-1657**) no asociados en la literatura a ningún haplotipo de HLA II (Figura 15).

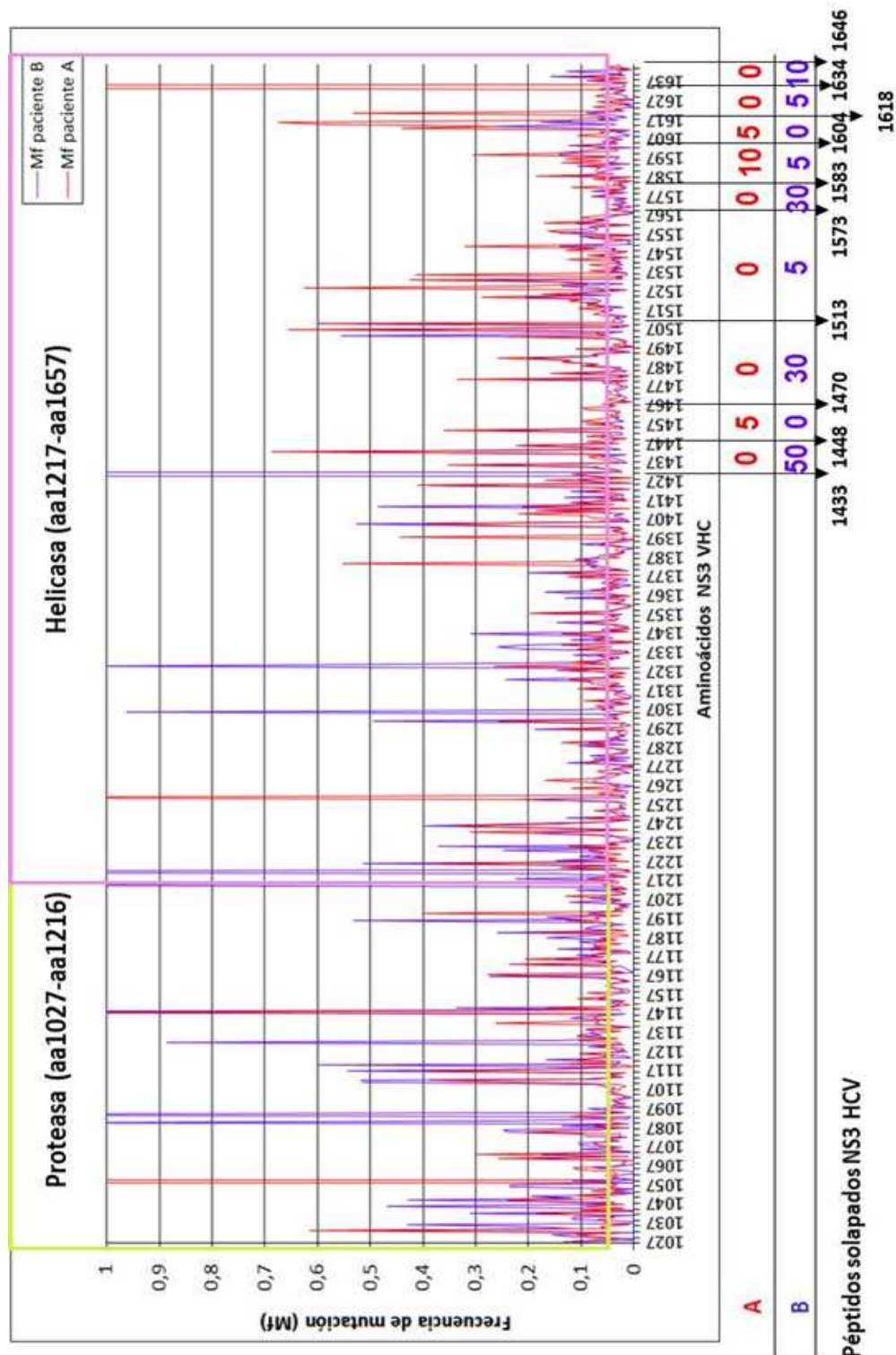


Figura 15. Frecuencia de mutación de aminoácidos del paciente A representado en rojo y el paciente B en azul en el eje de ordenadas del gráfico con respecto a las posiciones de NS3 del VHC en el eje de abscisas. En la parte inferior se indican los valores de ELISPOT en donde se produjo una mayor respuesta de CD4 antígeno- específica frente a epítomos de NS3 helicasa.

Los datos de complejidad de la población viral obtenidos por secuenciación masiva indicaron que el paciente A presentó mayor cambio de aa justo en la zona carboxi-terminal de la helicasa y en cambio la respuesta inmune fue inexistente, observándose todo lo contrario en el paciente B (resolución espontánea de la infección) que tenía una buena respuesta inmune, pero pocos cambios de aa (Figura 15).

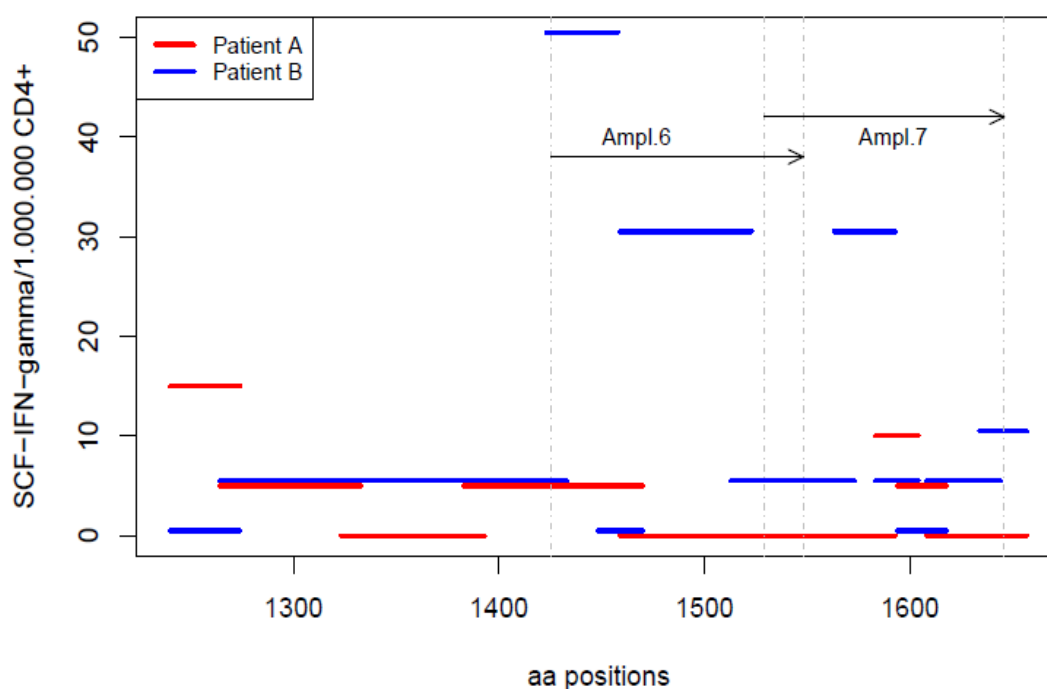


Figura 16. Resultados de ELISPOT de las células CD4 productoras de IFN-γ de los pacientes A y B frente al total de 13 péptidos que cubren todo el gen de NS3. Los valores indican números de "spots" de células CD4 productoras de IFN-γ frente a péptidos de NS3 helicasa.

La respuesta inmune en el paciente agudo B fue multiespecífica contra la región carboxi-terminal (la mitad de la helicasa), mientras que el paciente crónico A fue casi inexistente (Figura 16).

5.3.7 Densidad de mutación poblacional (Montserrat plots)

Una manera de visualizar la diversidad poblacional respecto al haplotipo más representativo, consistió en representar plots de densidad mutacional. Tomando la secuencia más representada como haplotipo de referencia, se contabilizó el número de mutantes de 1 posición, 2, 3 y así sucesivamente, se calculó su frecuencia y se representó en plots de densidad de mutación poblacional (bautizados como Montserrat plots por su similitud a las montañas con este nombre). Cuantas más mutaciones de diferencia se detecten, más diversa será la población. Fue posible la representación de la distribución esperada de simples, dobles, triples, etc de mutantes en la población genómica siguiendo una distribución de Poisson en base al número promedio de mutaciones determinadas experimentalmente para dicha población genómica.

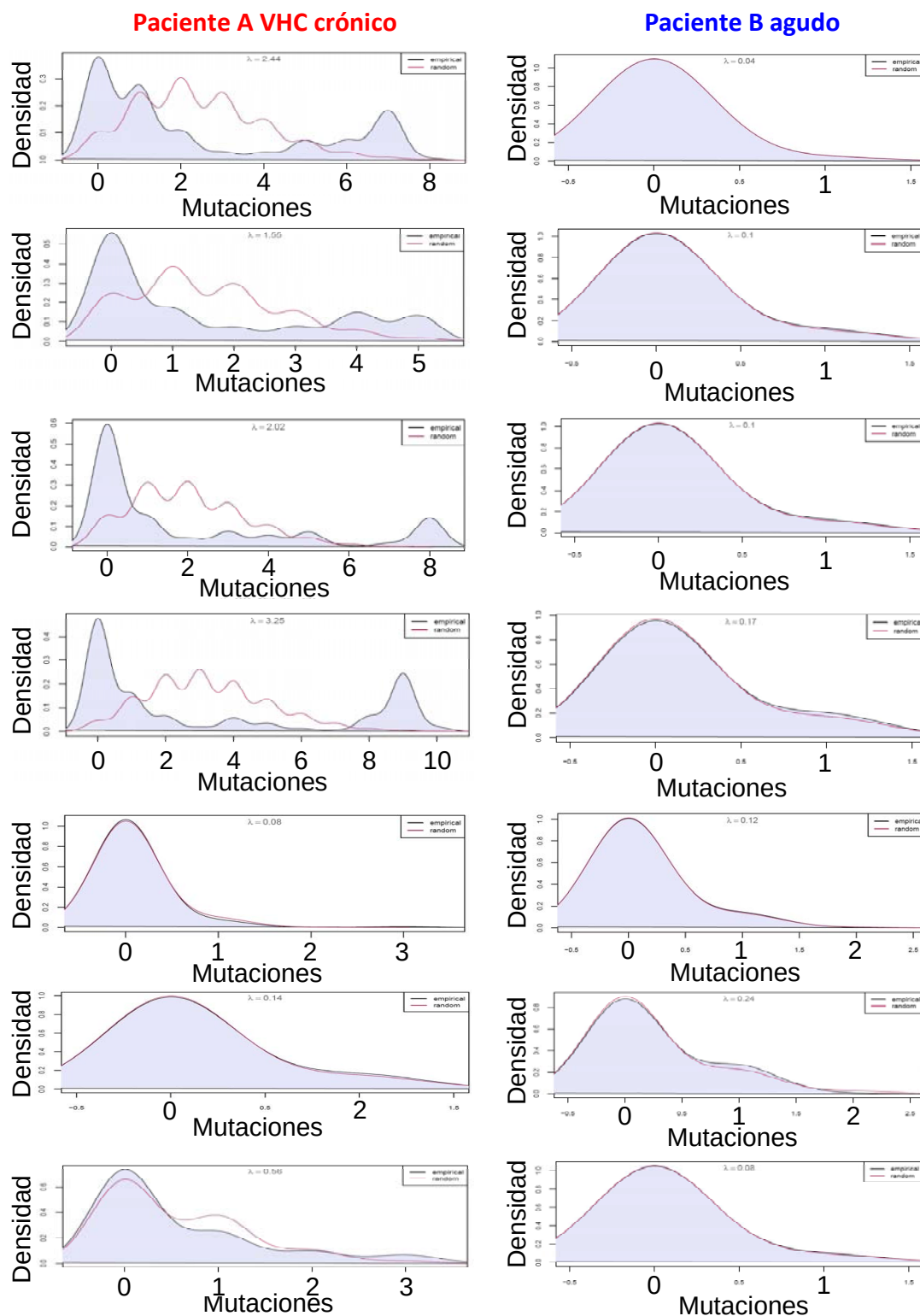


Figura 7. Distribución de densidad de mutación respecto al haplotipo dominante en los 7 amplicones estudiados. En los diagramas de la izquierda se representa como se distribuye la población viral del paciente A y en los de la derecha el paciente B.

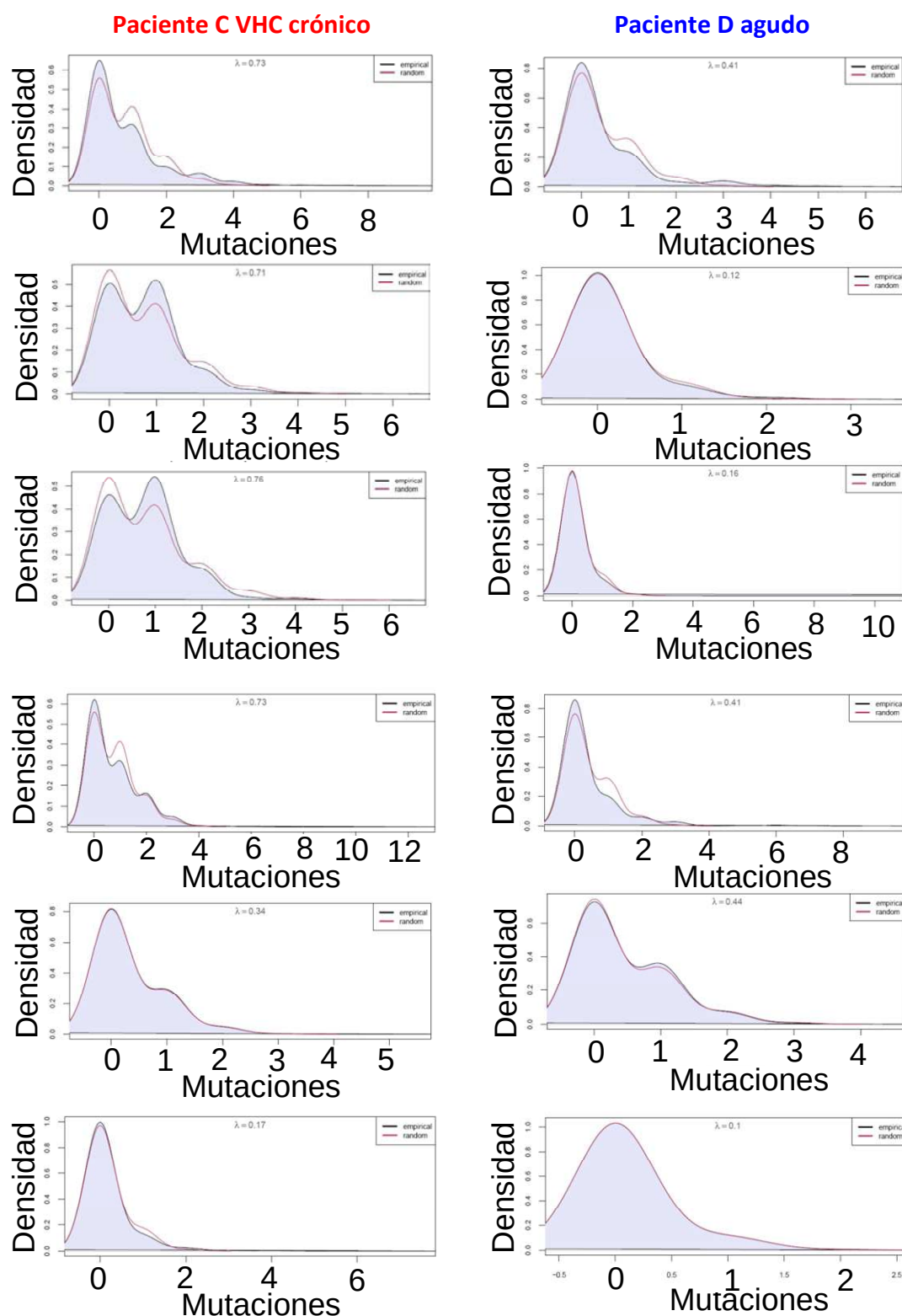


Figura 8. Diagramas de densidad de mutación de los pacientes C y D que corresponden a los amplicones del 2-7, el 1 no se incluyó porque la longitud de los “reads” fue corta y no cubría toda la región del amplicón 1.

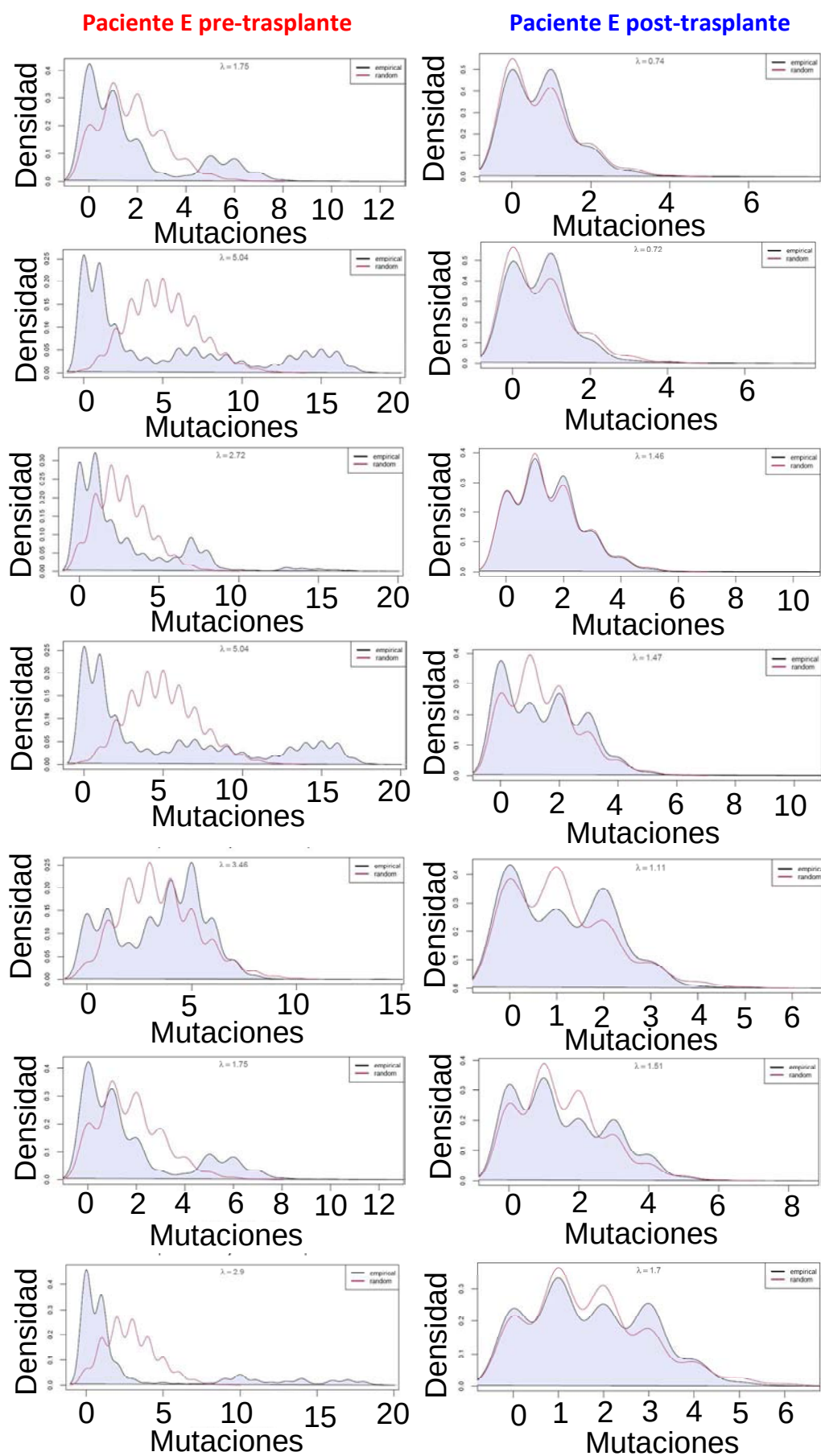


Figura 9. Diagramas de densidad poblacional del paciente E antes y después del trasplante.

La frecuencia de mutantes del paciente agudo B (Figura 7) en los 7 amplicones siguió un patrón de distribución idéntico al esperado si se tratase de una distribución debida al azar. Esto sugiere que en el paciente B las mutaciones aparecieron de forma aleatoria, tal como indica el estudio de la relación dN/dS. El paciente crónico A también presentó una distribución al azar en los amplicones 5 y 6. En cambio, en los amplicones del 1 al 4, y ligeramente en el 7, esta distribución difirió del simple azar. Este resultado sugiere que la población viral podría estar sometida a presión selectiva en estas regiones. Sorprendentemente, en los amplicones del 1 al 4, la población viral presentó una estructura bimodal donde se distinguieron 2 subpoblaciones, la población alrededor de la secuencia mayoritaria y una segunda población alrededor de secuencias con 3 a 10 mutaciones (según haplotipo).

En el caso de la 2ª pareja de pacientes (C y D), la curva real y la curva teórica se superpusieron, sugiriendo en ambos casos una situación de acumulación de mutaciones al azar, con ligeras variaciones en el paciente crónico (Figura 8).

En el paciente cirrótico E pre-transplante, la curva real fue diferente a la de azar en los 7 amplicones estudiados, indicando que la población viral está lejos del azar (Figura 9). Es interesante observar la enorme heterogeneidad de la población que apareció en la situación de cirrosis antes del transplante, presentando incluso una distribución trimodal en los amplicones 2 y 4. Sería interesante comprobar con un número mayor de pacientes cirróticos si esta distribución se asocia a daño hepático.

Análisis comparativo entre parejas de pacientes relacionados

5.3.8 *Identidad poblacional mediante coeficiente de Bhattacharyya*

El coeficiente de Bhattacharyya (Tabla 8) permite representar la identidad (similitud) entre 2 poblaciones de virus. Este coeficiente es igual a 1 cuando las poblaciones virales son idénticas y es igual a 0, en el caso de que no exista ningún haplotipo idéntico.

Tabla 8. Valores de coeficiente de Bhattacharyya en los aislados de las parejas de pacientes A-B, C-D y E pre-transplante y E post-transplante.

Coeficiente de Bhattacharyya			
Pacientes			
Amplicones	A y B	C y D	E-pre y E-post
1	0,00000	*	0,278630
2	0,00000	0,81404	0,145669
3	0,00000	0,67054	0,087599
4	0,02812	0,68253	0,09290
5	0,00000	0,14642	0,232259
6	0,78055	0,21095	0,518537
7	0,07479	0.050881	0,625183

La pareja de aislados con mayor identidad poblacional fue la de los pacientes C y D (con valores cercanos a 1, excepto en los amplicones 6 y 7), seguido de E-pre y E-post, y finalmente, la pareja de pacientes A y B (transmisión del VHC de tipo cuello de botella genético-estudio 1) fue la que presentó una divergencia mayor, excepto en una subregión de la helicasa (amplicón 6) donde ambas poblaciones coincidieron.

El coeficiente de Battacharyya también se representó por sliding window, donde se puede observar que en los pacientes C y D (Figura 10b) es donde existió una identidad poblacional mayor, mientras que la pareja de pacientes A y B fue donde se produjeron menos coincidencias, excepto en el amplicón 6 (Figura 10a).

El coeficiente de Battacharyya también se representó gráficamente (Figura 10).

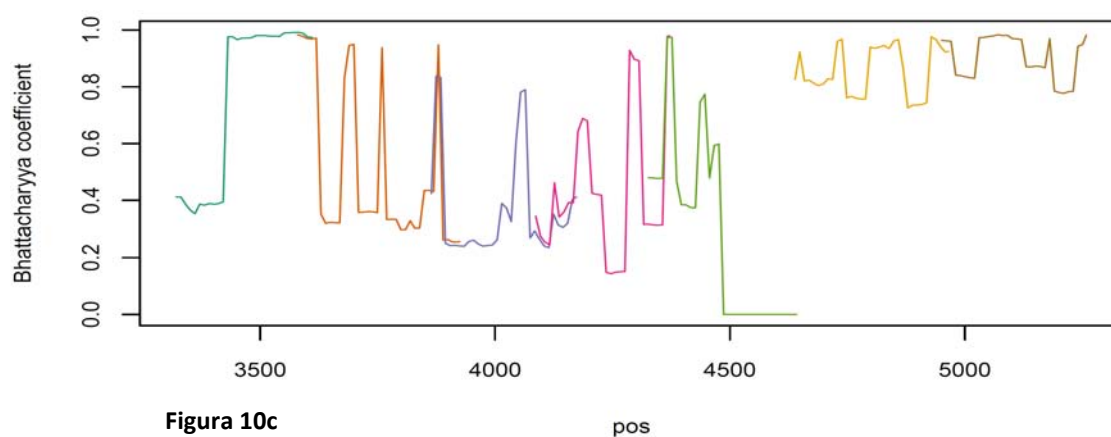
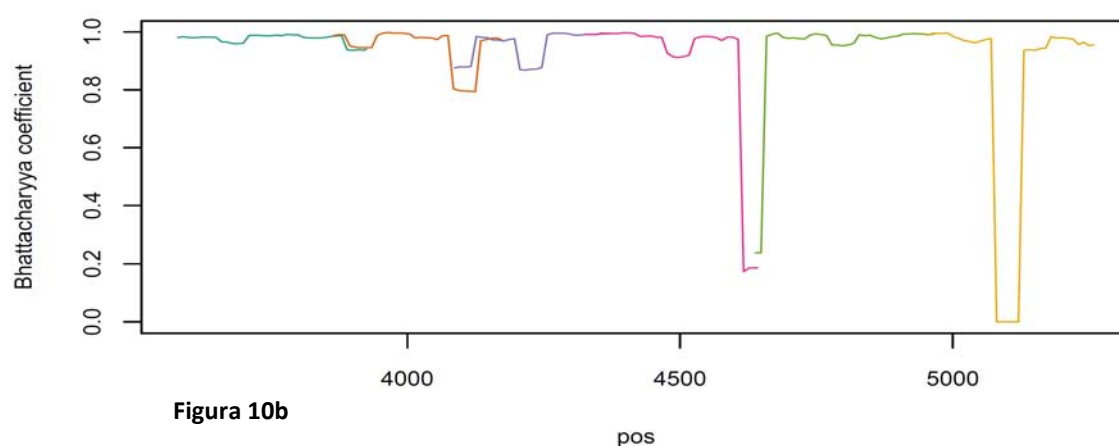
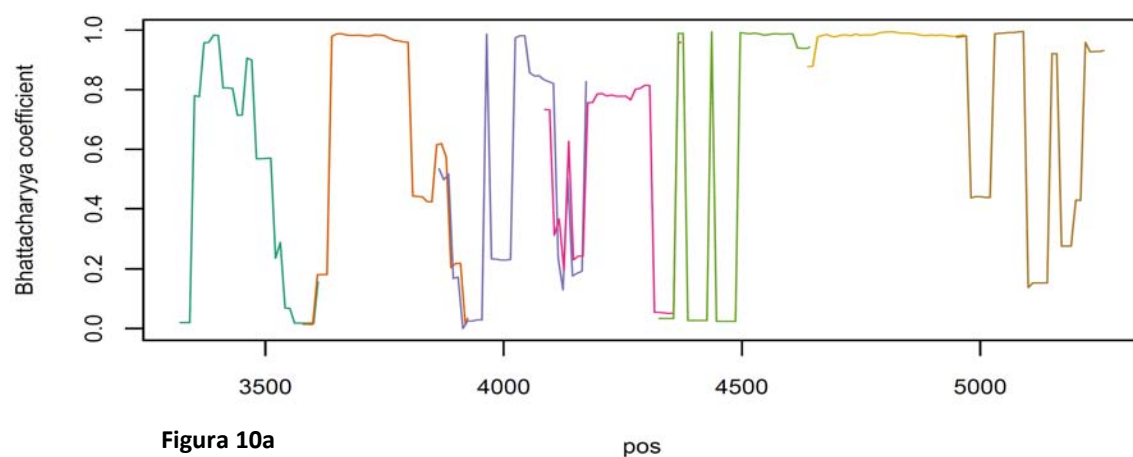


Figura 10. Coeficiente de Battacharyya como medida de identidad poblacional entre los pacientes A y B (10 a), C y D (10b) y Epre y Epost (10c) por sliding window de 50 nucleótidos en pasos de 10 nucleótidos. Las secuencias de cada par de pacientes se compararon una a una en ventanas de 50 nucleótidos.

5.3.9 *Distancia genética*

El hecho de que cada pareja de pacientes haya compartido inóculo, permitió comparar cuanto se parecen las poblaciones después de la transmisión. Para ello se realizó el cálculo de la distancia genética (Da) de los 7 amplicones, en la cual, cuanto más cercano el valor de Da a 0, más se parecen las 2 poblaciones virales comparadas.

Tabla 9. Valores de distancia genética

Distancia genética (Da)			
Pacientes			
Amplicones	A y B	C y D	E-pre y E-post
1	6,60703105	1,02802194	3,8692578
2	5,55207707	0,01608232	7,7724927
3	5,2322938	0,11962913	10,7526256
4	2,48798052	0,12778528	5,9210072
5	3,00292196	0,95890033	4,4039158
6	0,05108796	0,92084094	0,6914808
7	2,29067187	1,01017173	0,3886573

En el caso de la pareja de pacientes A y B, todos los amplicones (excepto el 6) presentaron valores de Da superiores a 2 (Tabla 9). En el caso de los pacientes C y D, la población viral prácticamente no varió, con valores de Da en los 7 amplicones no superiores a 1. Con respecto a los 2 aislados del paciente E, se produjo una gran divergencia de la población vírica con valores entre 3 y 10, excepto en los amplicones 6 y 7.

5.3.10 Árboles filogenéticos

El estudio de pacientes que han compartido inóculo ofreció la oportunidad de analizar la relación filogenética entre los virus (secuencias clonales) identificados por secuenciación masiva de los aislados de cada paciente. El análisis filogenético incluyó secuencias que estuvieran representadas por más de un 0,5% de lecturas (“reads”), y se incluyeron las secuencias consenso de aislados de dichos pacientes (Figuras 11, 12 y 13).

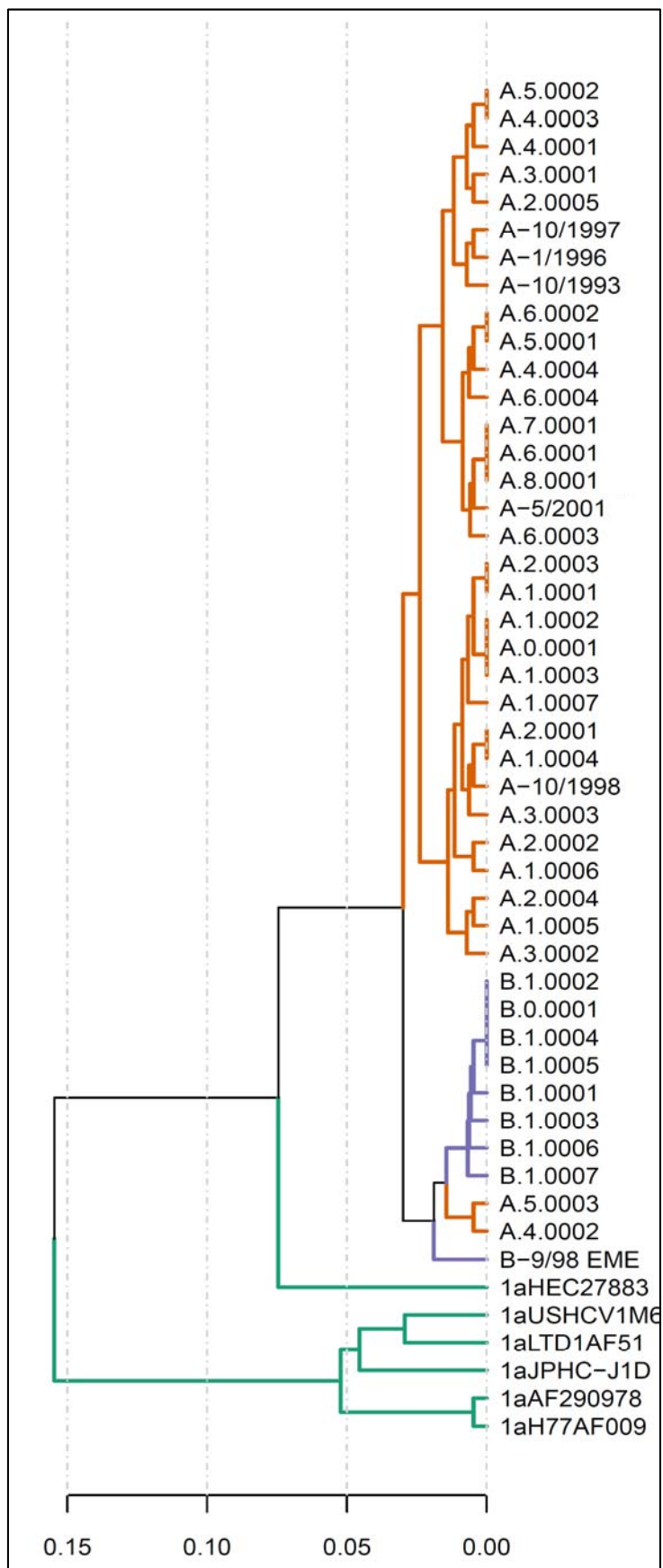
Árboles filogenéticos pacientes A y B

En el caso de los pacientes A y B, el resultado más interesante del análisis filogenético fue que 2 haplotipos del paciente A con 4 y 5 mutaciones (A.4.002 y A.5.003) en el amplicón 1 (Figura 11a), un haplotipo con 5 mutaciones (A.5.003) en el amplicón 3 (Figura 11c) y un haplotipo con 3 mutaciones (A.3.003) en el amplicón 7 (11g), se agruparon con la secuencia mayoritaria y el resto de haplotipos virales del paciente B, en cada uno de los amplicones nombrados. Este resultado sugiere que la infección en B pudo iniciarse por un número reducido de virus del paciente A, tal como se expuso en el trabajo del estudio nº 1 (página 22).

Figura 11. Árboles filogenéticos UPGMA de amplicones 1 (11 a), 2 (11 b), 3 (11 c), 4 (11 d), 5 (11e), 6 (11 f) y 7 (11 g) de los pacientes A y B (Ver páginas 72-78). Sólo secuencias con más de un 0,5% de “reads” están representadas. Las secuencias representadas en naranja corresponden al paciente A, las secuencias representadas en azul a la paciente B, y en verde secuencias obtenida de la base de datos GenBank. Cada haplotipo se enumeró como Z.n.mmmm, donde Z era la identificación del paciente, n se refería al número de mutaciones respecto al haplotipo dominante y mmmm era el número de orden del haplotipo con n mutaciones, donde el orden de la frecuencia respectiva fue decreciente (por ejemplo A.6.004 es el 4º haplotipo más abundante del paciente A con 6 mutaciones respecto a la secuencia más representada). Cada secuencia consenso se enumeró como Z-mm/yyyy, en el que mm= mostró el mes e yyyy= año (por ejemplo A-5/2001= un aislado del paciente A aislado en mayo de 2001). Los círculos en negro (amplicones 1, 3 y 7) engloban las

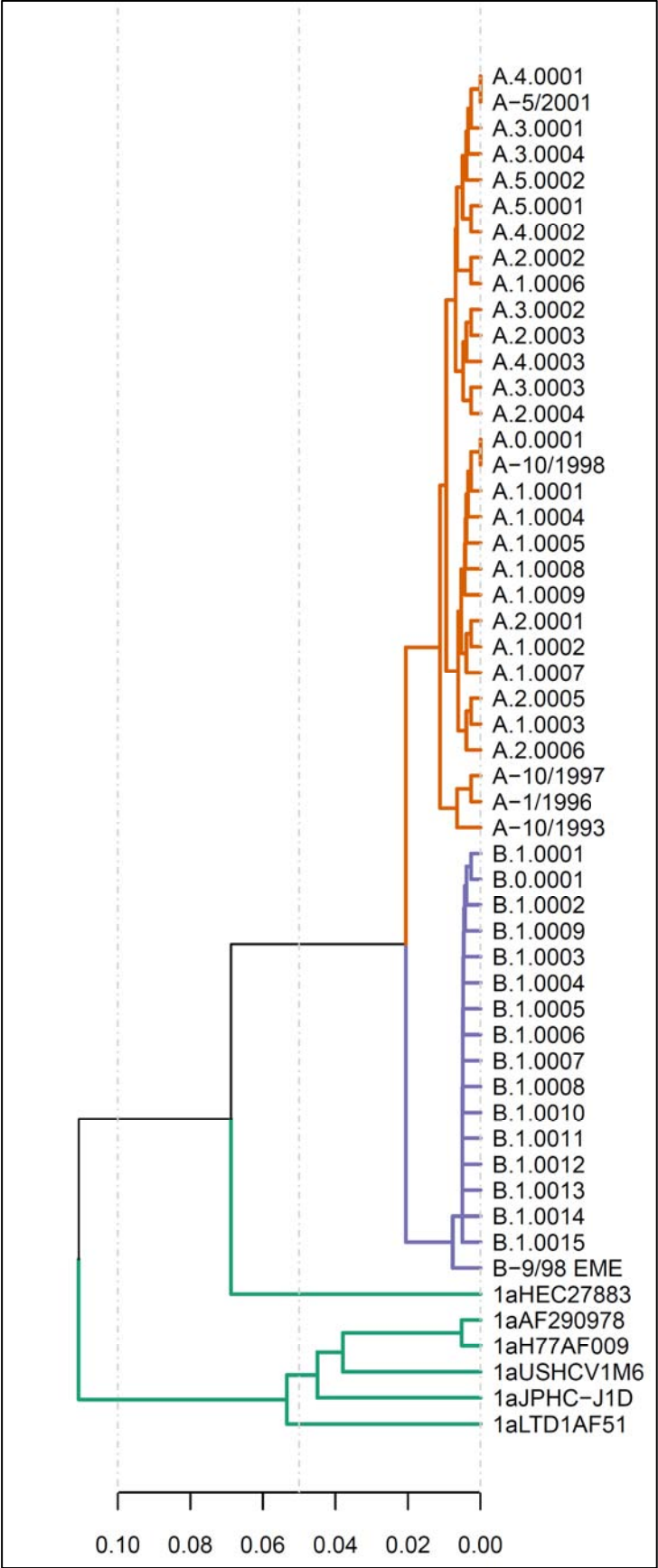
AMPLICÓN 1

A&B-3420:3634



AMPLICÓN 2

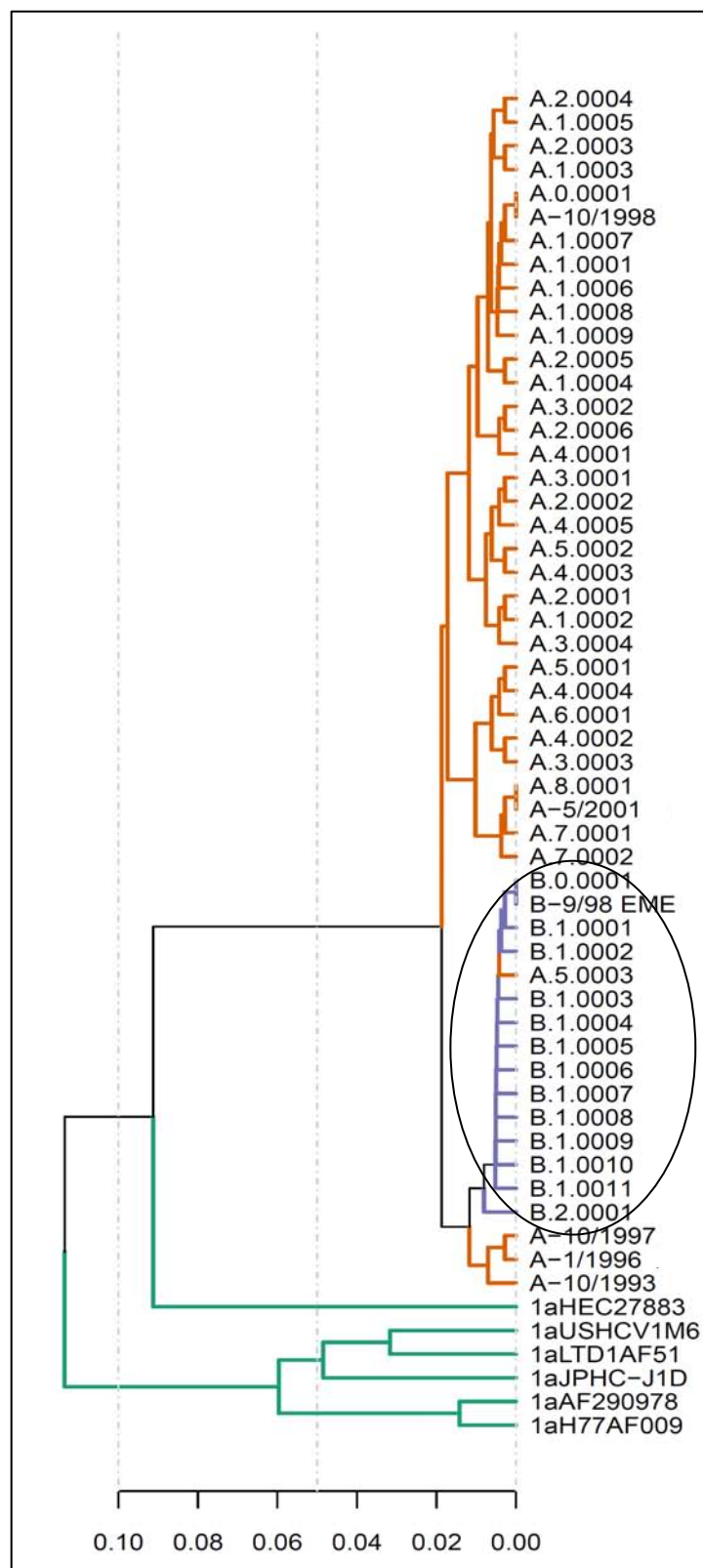
A&B-3555:3944



11 b

AMPLICÓN 3

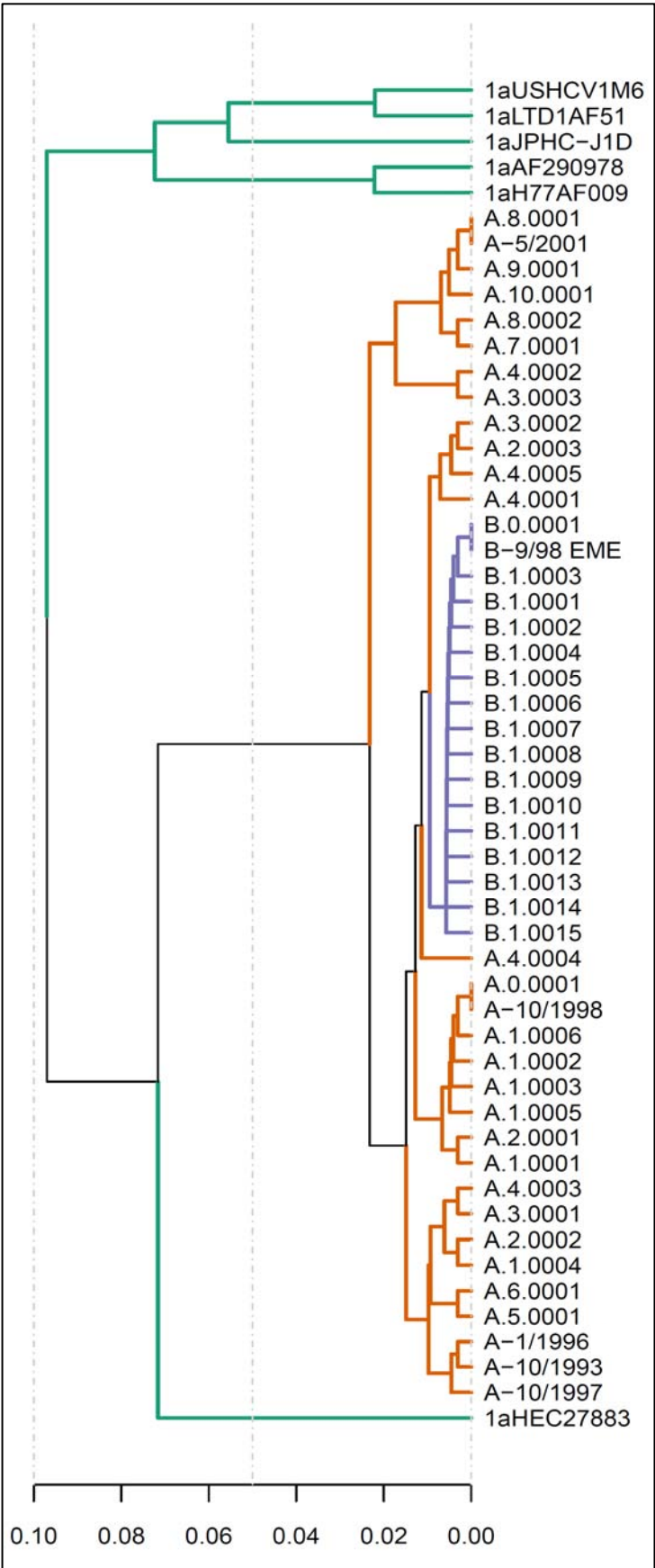
A&B-3840:4196

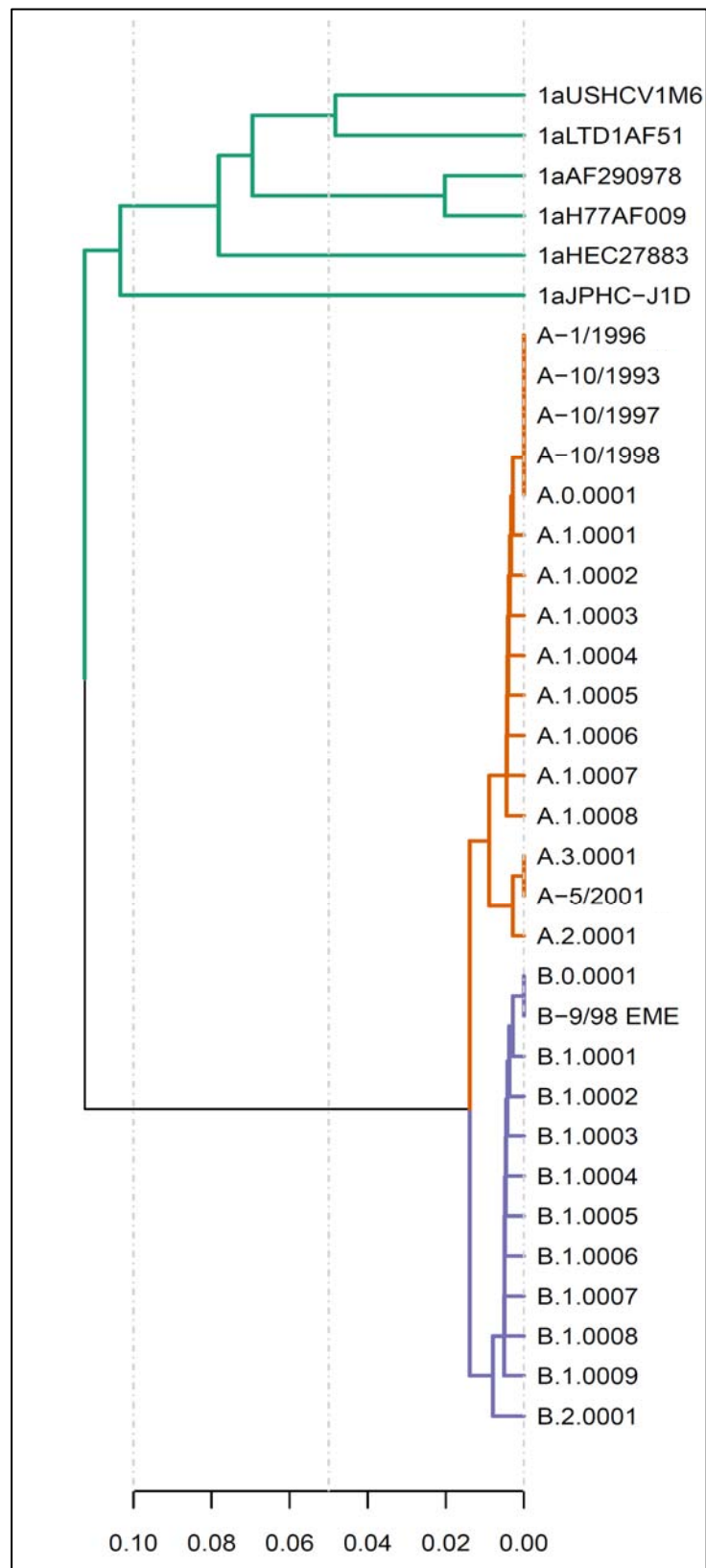


11 c

AMPLICÓN 4

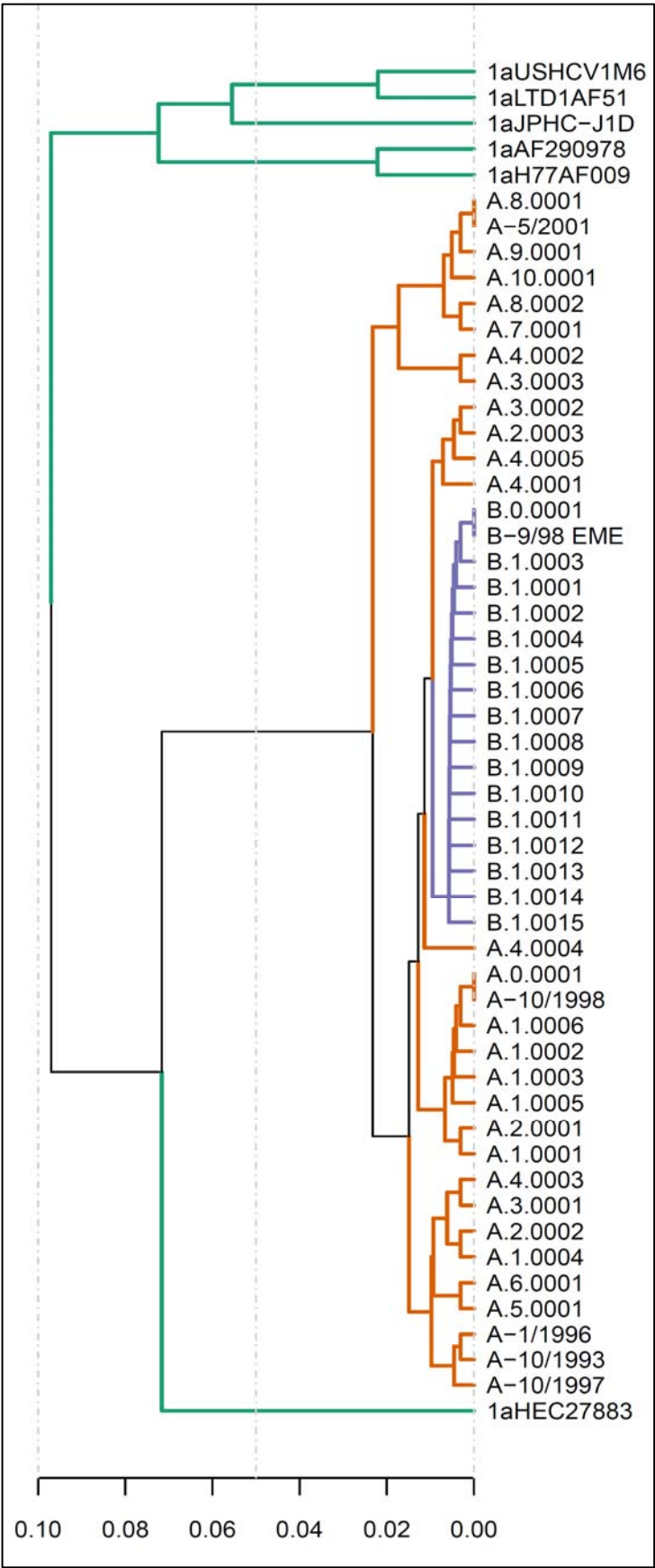
A&B-4062:4391

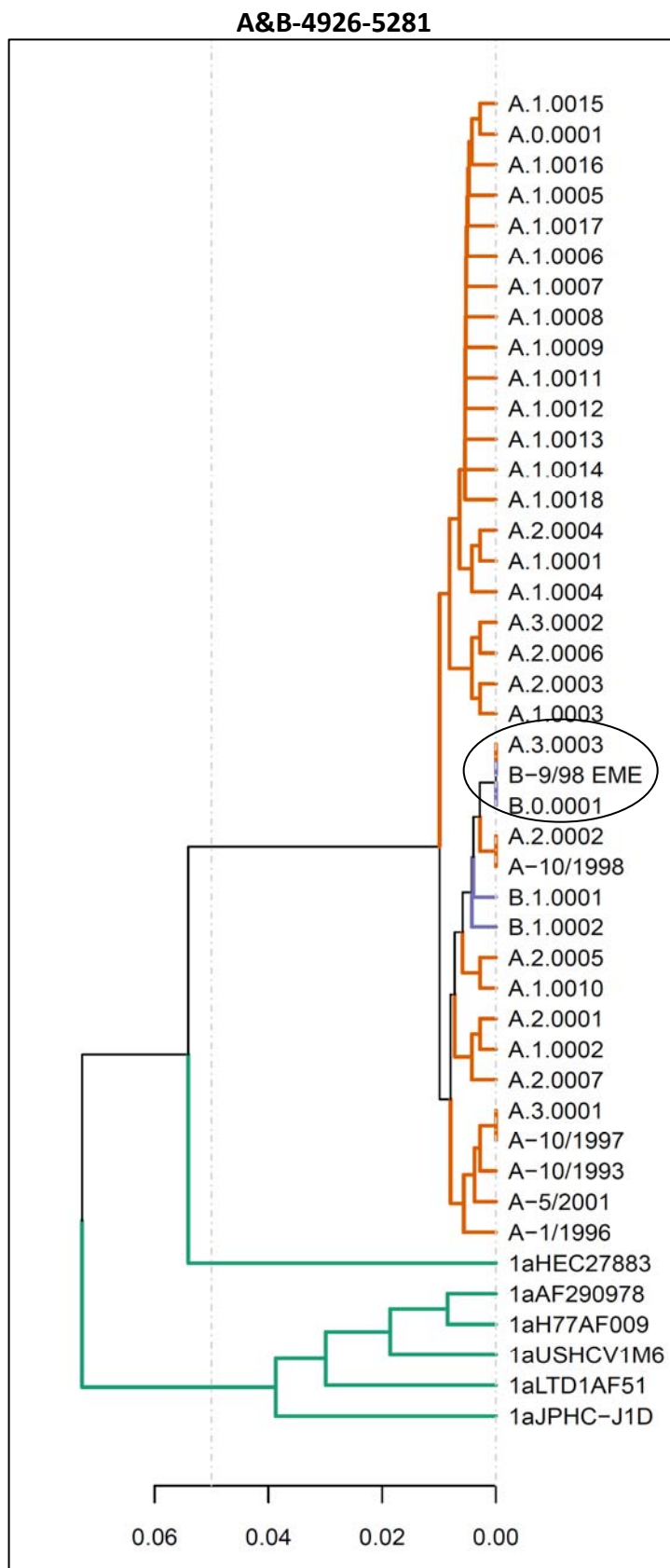


AMPLICÓN 5**A&B-4302:4661**

AMPLICÓN 6

A&B-4614:4986



AMPLICÓN 7

11 g

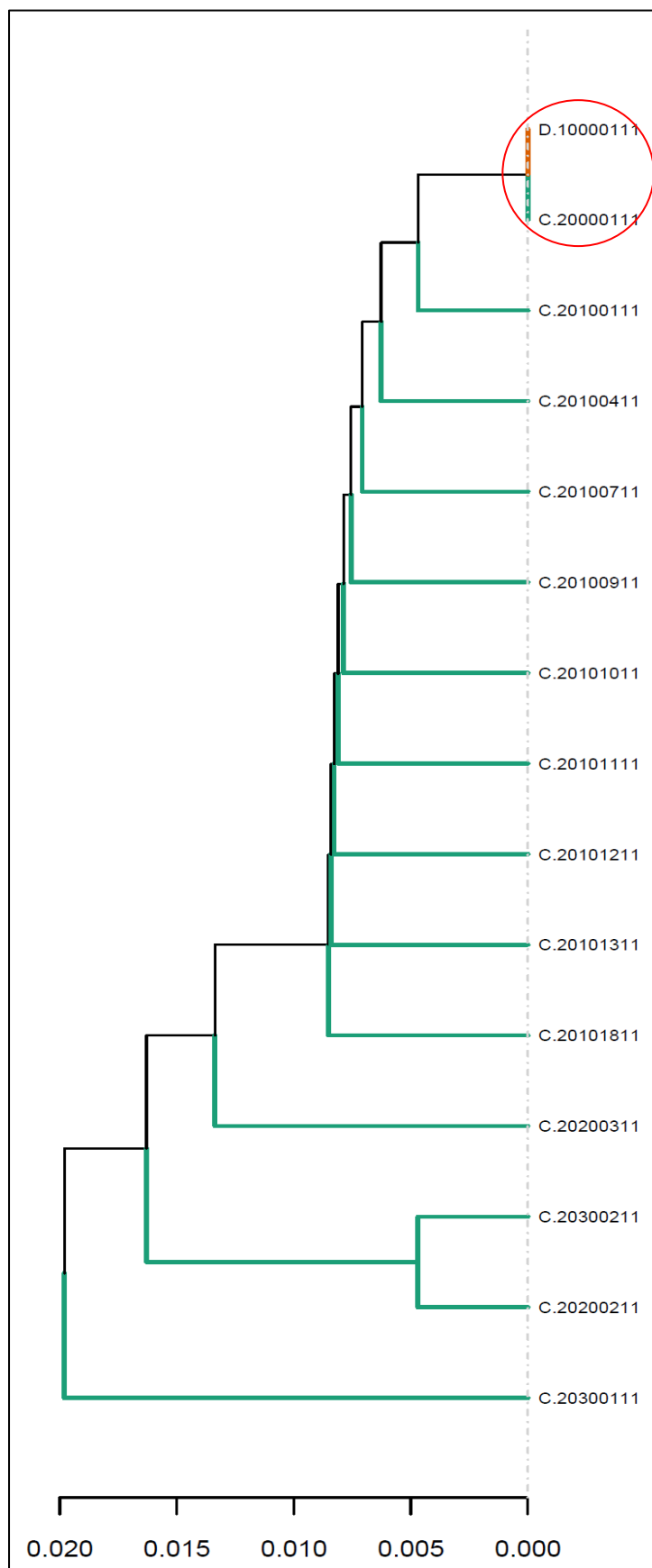
Árboles filogenéticos pacientes C y D

Con respecto al estudio de la filogenia de los pacientes C y D, en los amplicones del 1 al 4 coincidieron las secuencias dominantes. La población del VHC del paciente C que inició la infección en el paciente D se adaptó rápidamente al ambiente nuevo, sufriendo poca diversificación, ya que muchas secuencias con mutaciones de ambos pacientes coincidieron. Los 7 amplicones presentaron un perfil filogenético similar. Por tanto, o bien la infección se pudo iniciar a partir de la secuencia mayoritaria del paciente, o a partir de una representación importante de la población del paciente crónico (transmisión poblacional).

Figura 12. Árboles filogenéticos UPGMA de los amplicones 1 (**12a**), 2 (**12b**), 3 (**12c**), 4 (**12d**), 5 (**12e**), 6 (**12 f**) y 7 (**12g**). El círculo rojo en los amplicones 1, 2, 3 y 4 engloba las secuencias máster de ambos pacientes. (Representado en las páginas 80-86). La nomenclatura de cada haplotipo fue Z.nmmmmm, donde Z.n es la identificación del paciente, mm se refiere al número de mutaciones respecto al haplotipo dominante y mmm el nº de orden del haplotipo con n mutaciones, donde el orden de la frecuencia respectiva fue decreciente.

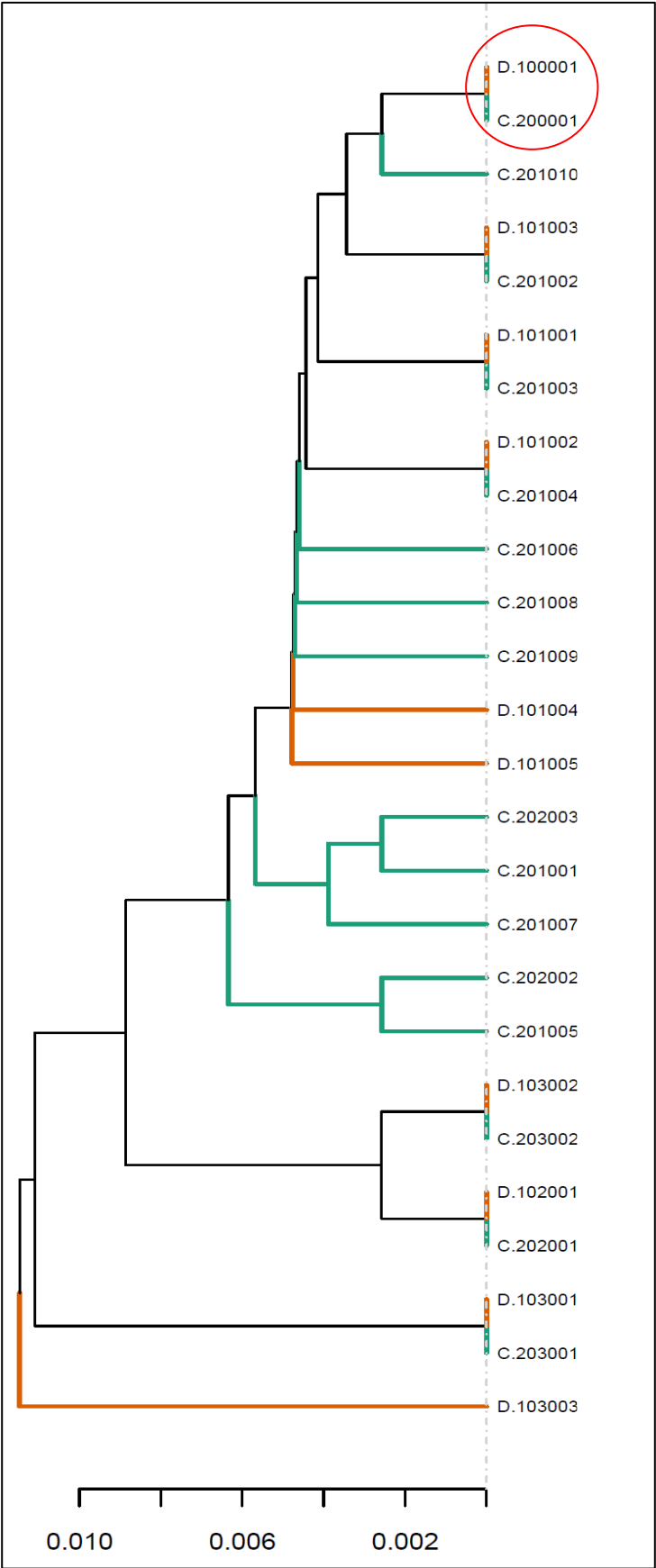
AMPLICIÓN 1

C&D-3420:3634



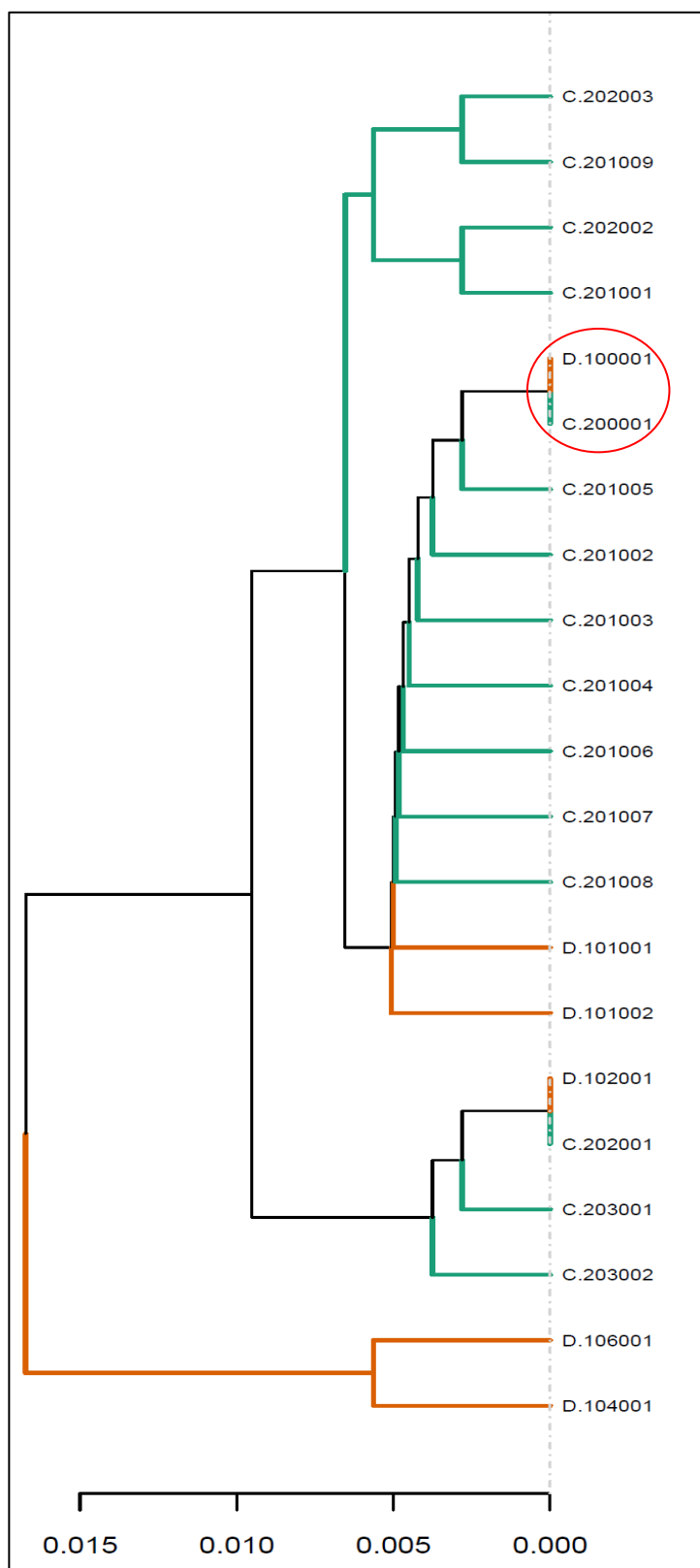
AMPLICÓN 2

C&D-3555:3944



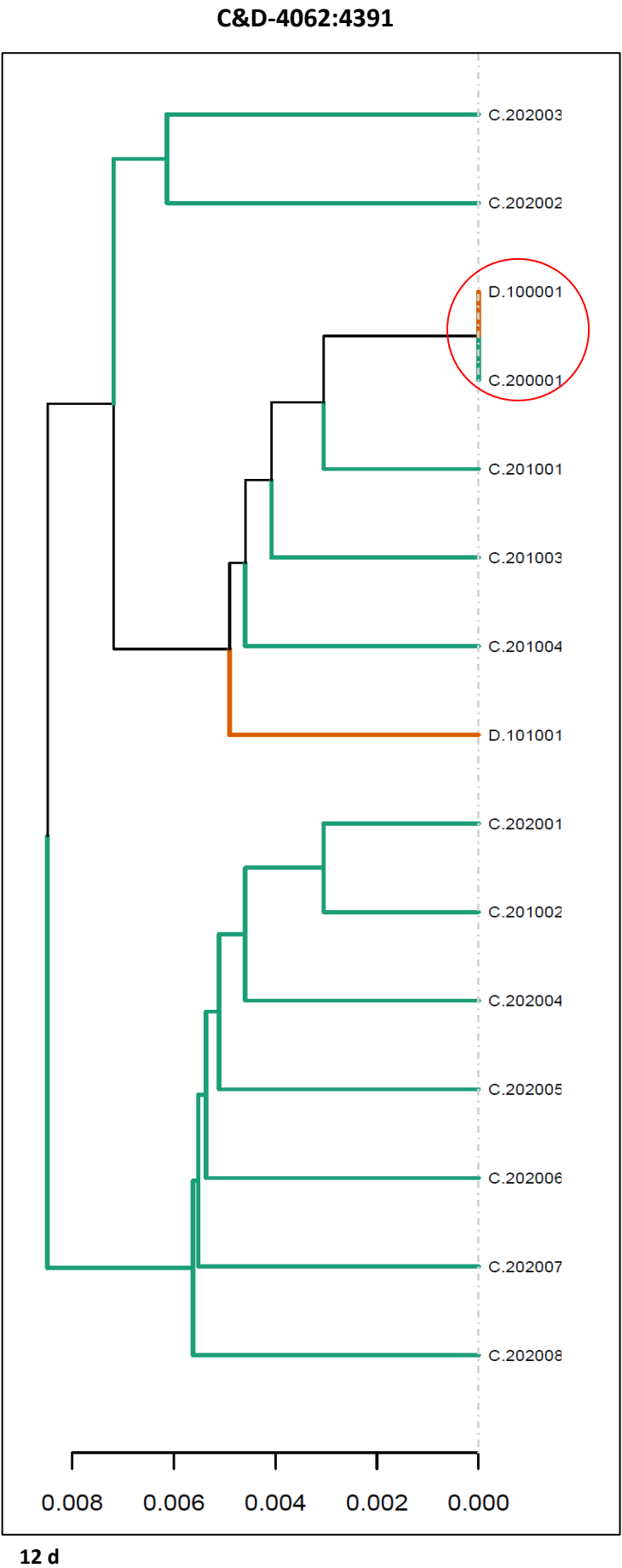
AMPLICÓN 3

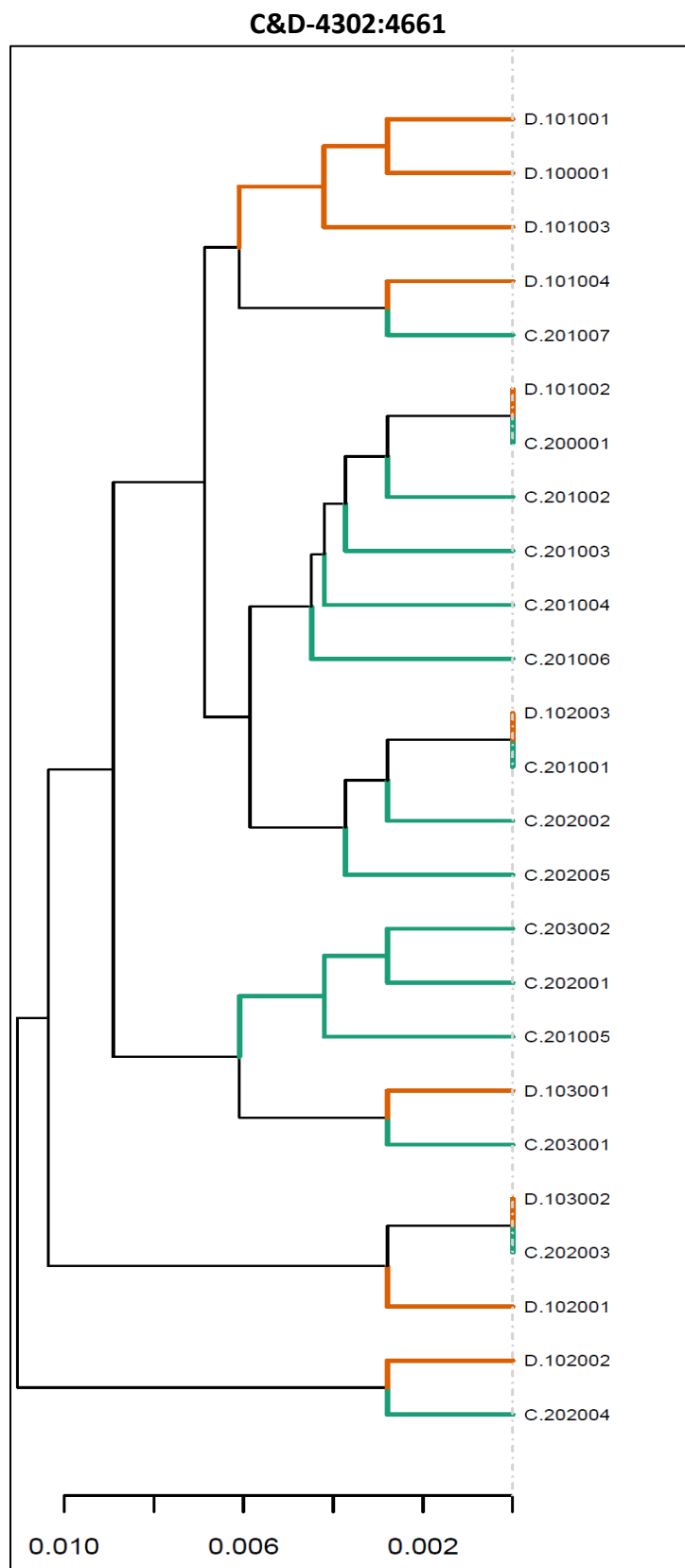
C&D-3840:4196



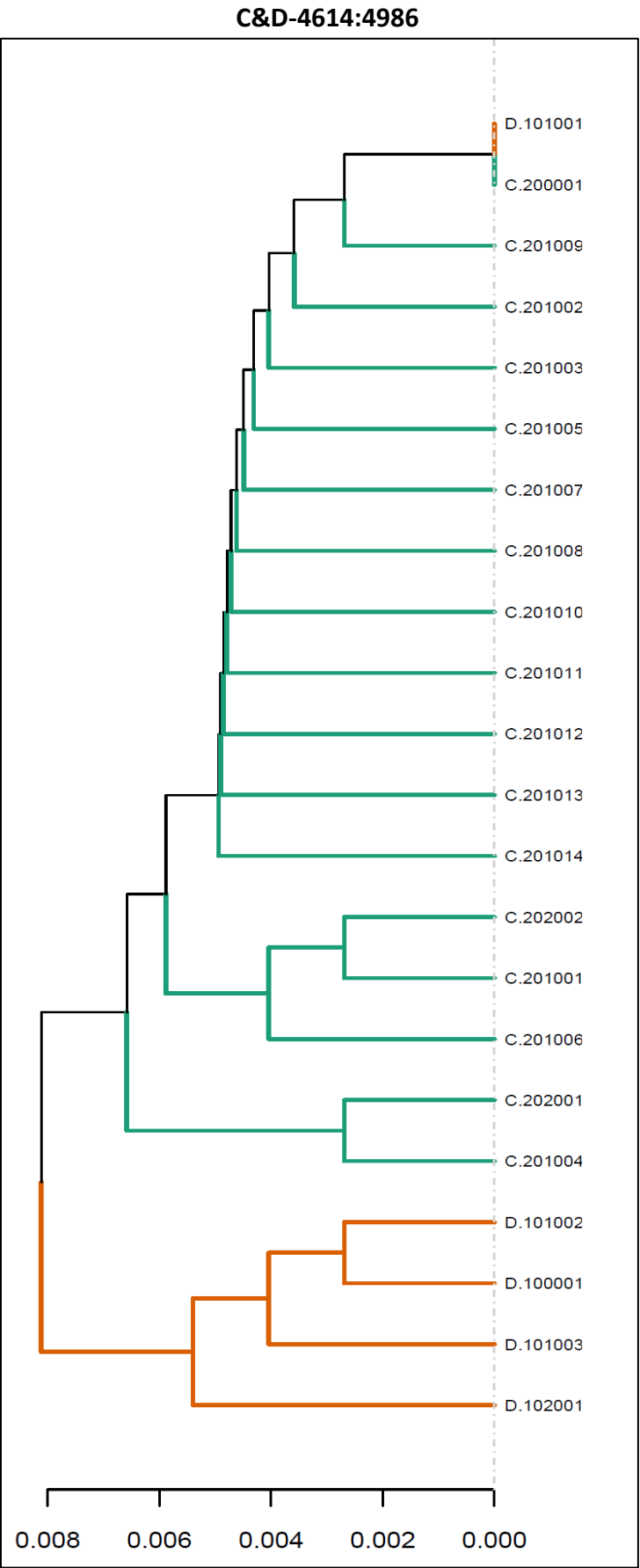
12 c

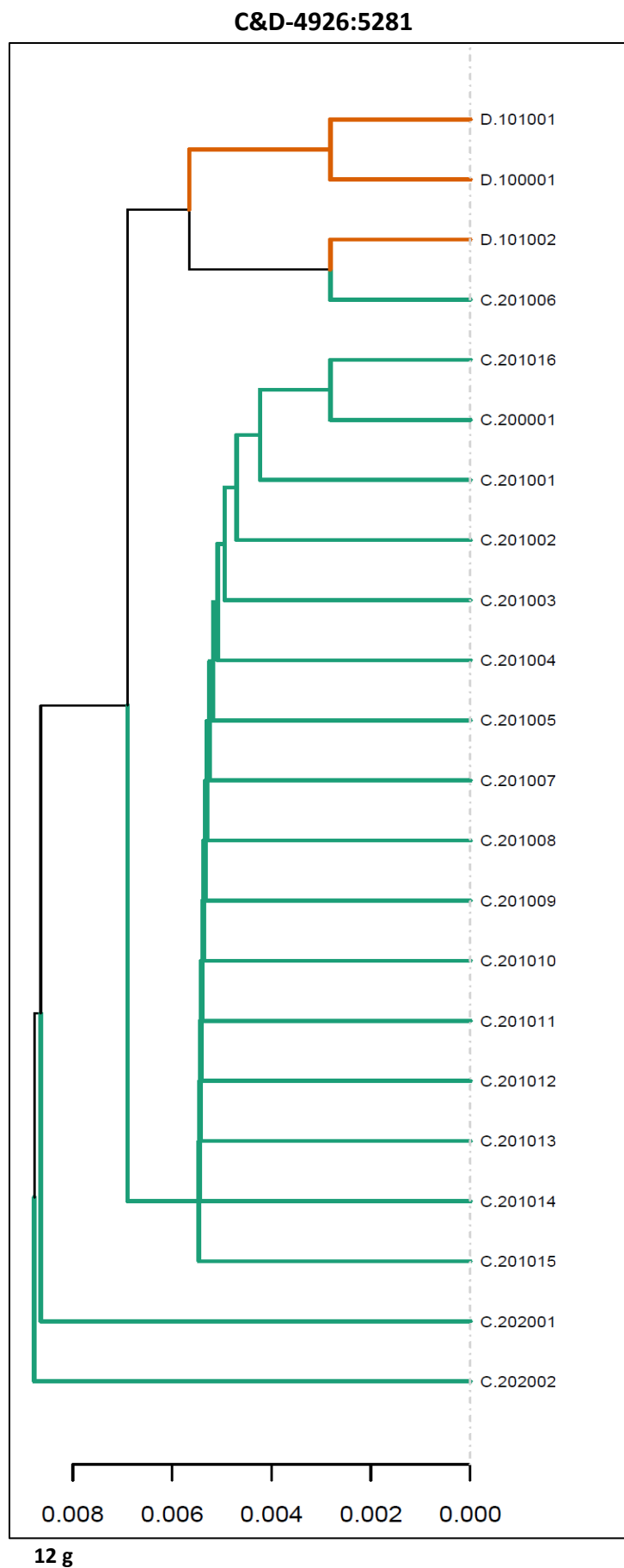
AMPLICÓN 4



AMPLICÓN 5

AMPLICÓN 6

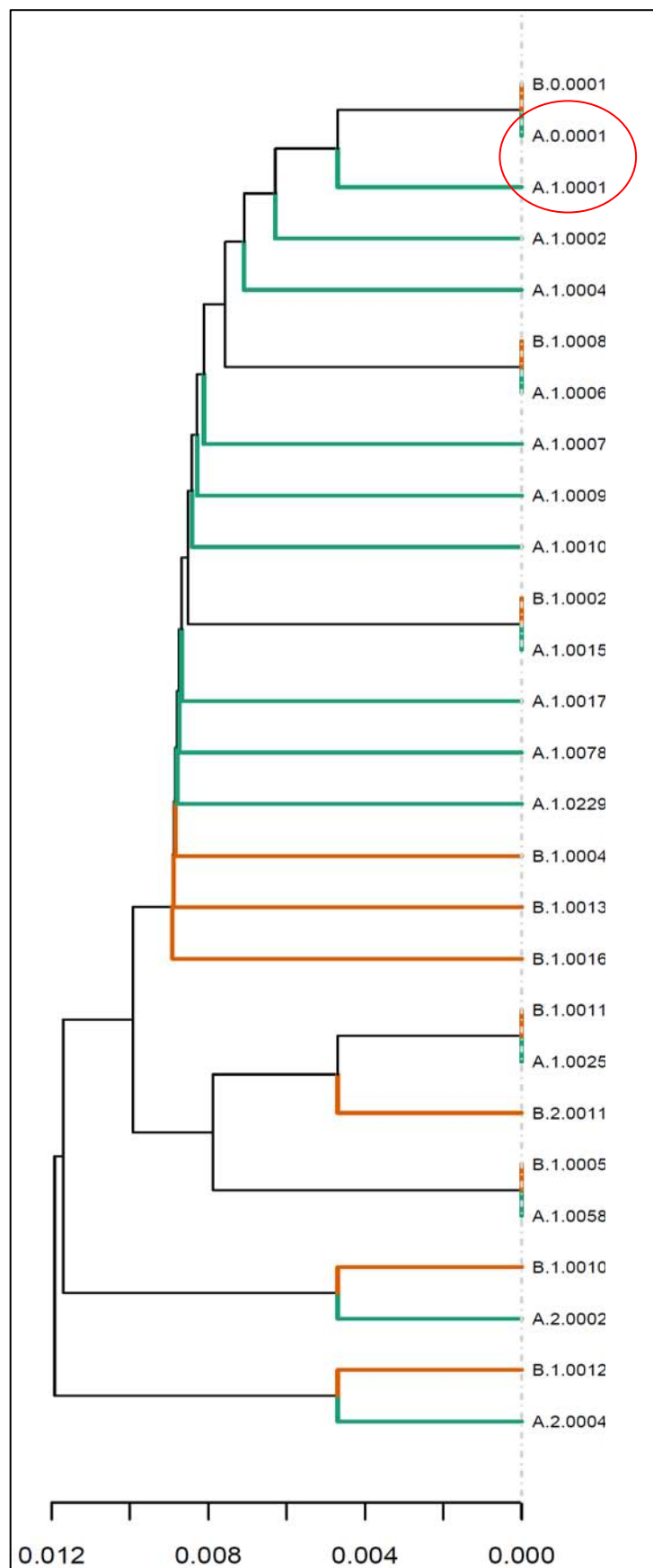


AMPLICÓN 7

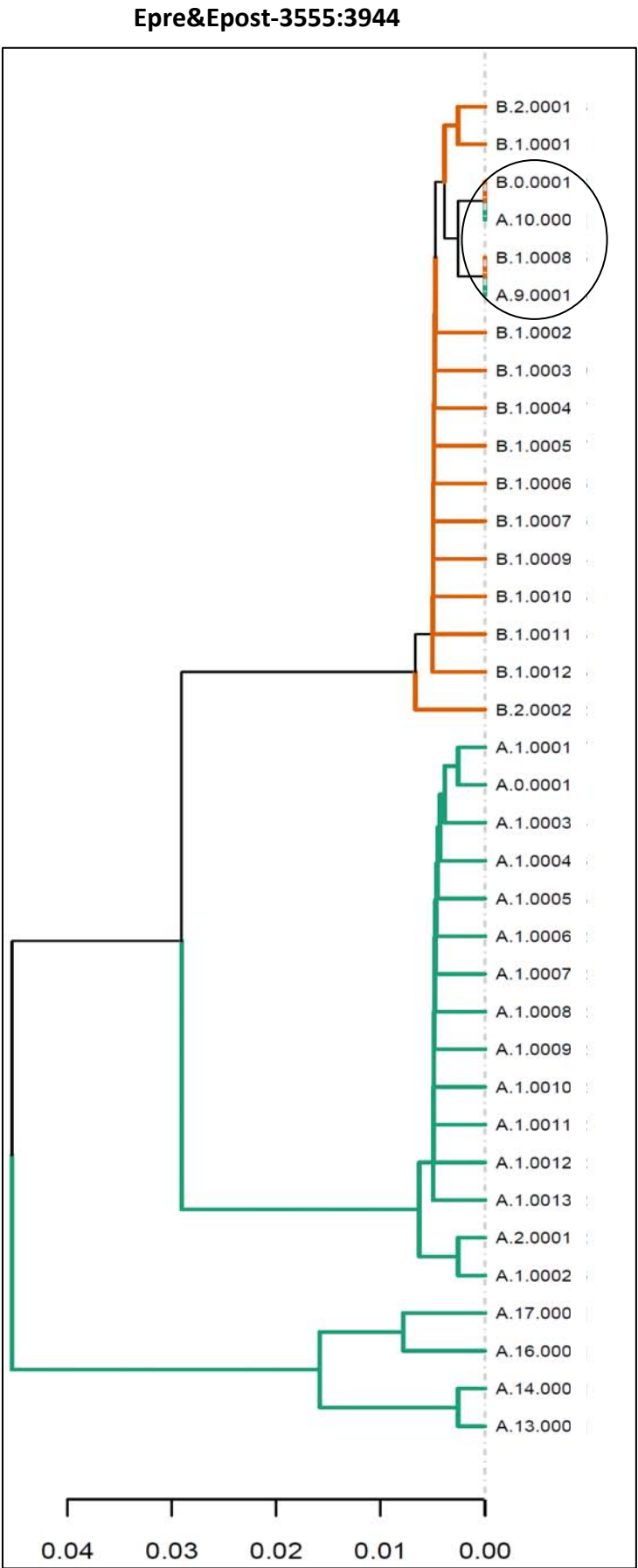
Árboles filogenéticos pacientes E-pre y E-post

En relación con el análisis filogenético de los pacientes Epre y Epost (Figura 13), mostró que en los amplicones 2, 3, 4 y 5 la población viral después del trasplante, provenía de secuencias del aislado del pre-trasplante con 9-10, 17-18, 8-9, 5-7 mutaciones respectivamente respecto a su secuencia mayoritaria. En los amplicones 1 y 6, las secuencias mayoritarias del Epre y Epost coincidieron, mezclándose ambas poblaciones, mientras que en el amplicón 7, las secuencias de Epre con 3-7 mutaciones fueron iguales a secuencias de la muestra de Epost.

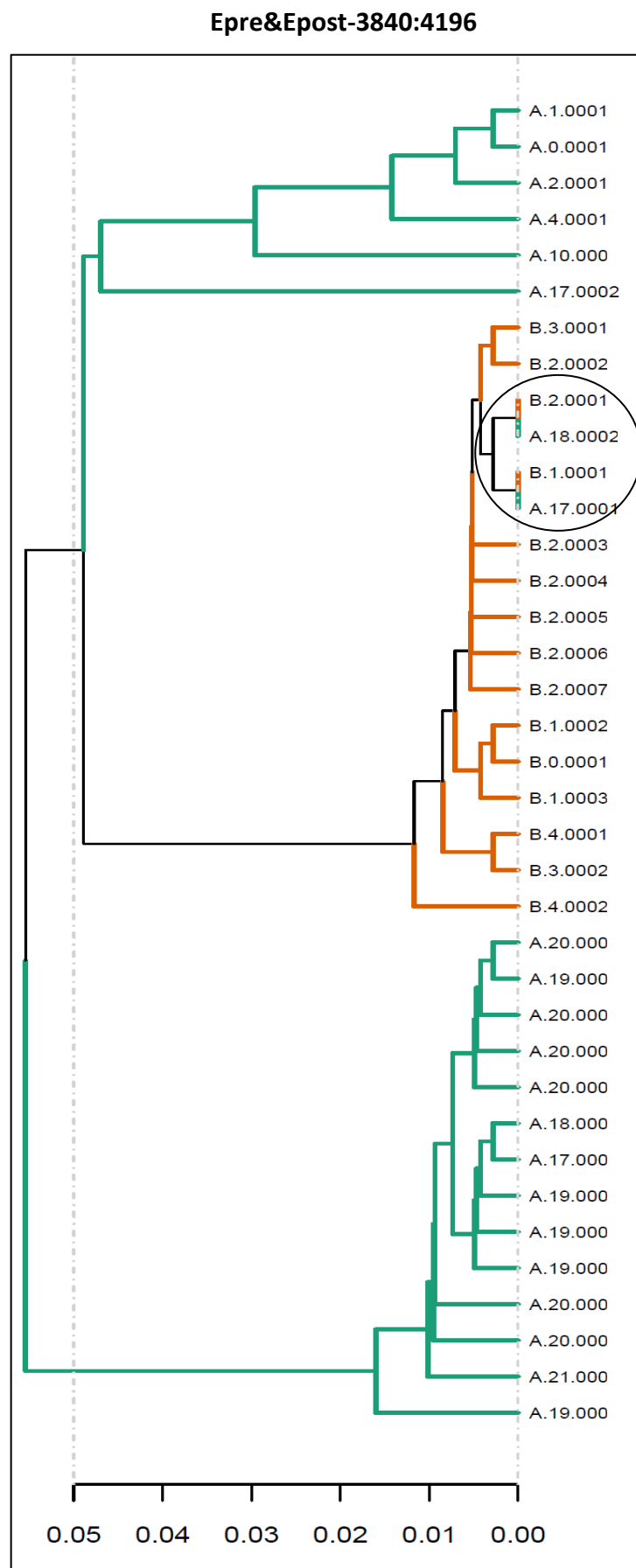
Figura 13. Árboles filogenéticos UPGMA de amplicones 1 (13 a), 2 (13b), 3 (13c), 4 (13d), 5 (13e), 6 (13f), y 7 (13g) del paciente E pre-trasplante y E post-trasplante. La nomenclatura es la misma que para los pacientes A y B. Los círculos en negro engloban las secuencias que están relacionadas (amplicones 2, 3 4 y 5). El círculo en rojo en los amplicones 1 y 6 engloba las secuencias máster de ambos pacientes.

AMPLICÓN 1**Epre&Epost-3420:3634**

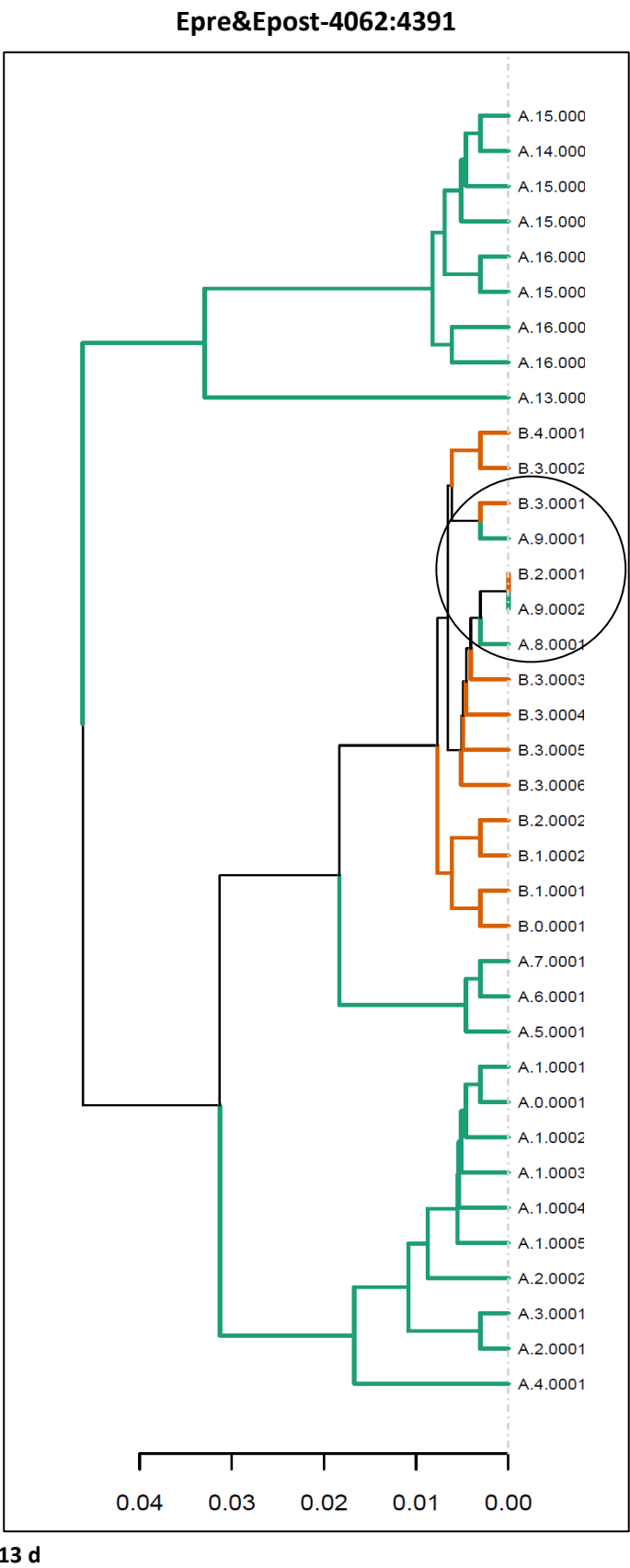
AMPLICÓN 2



13 b

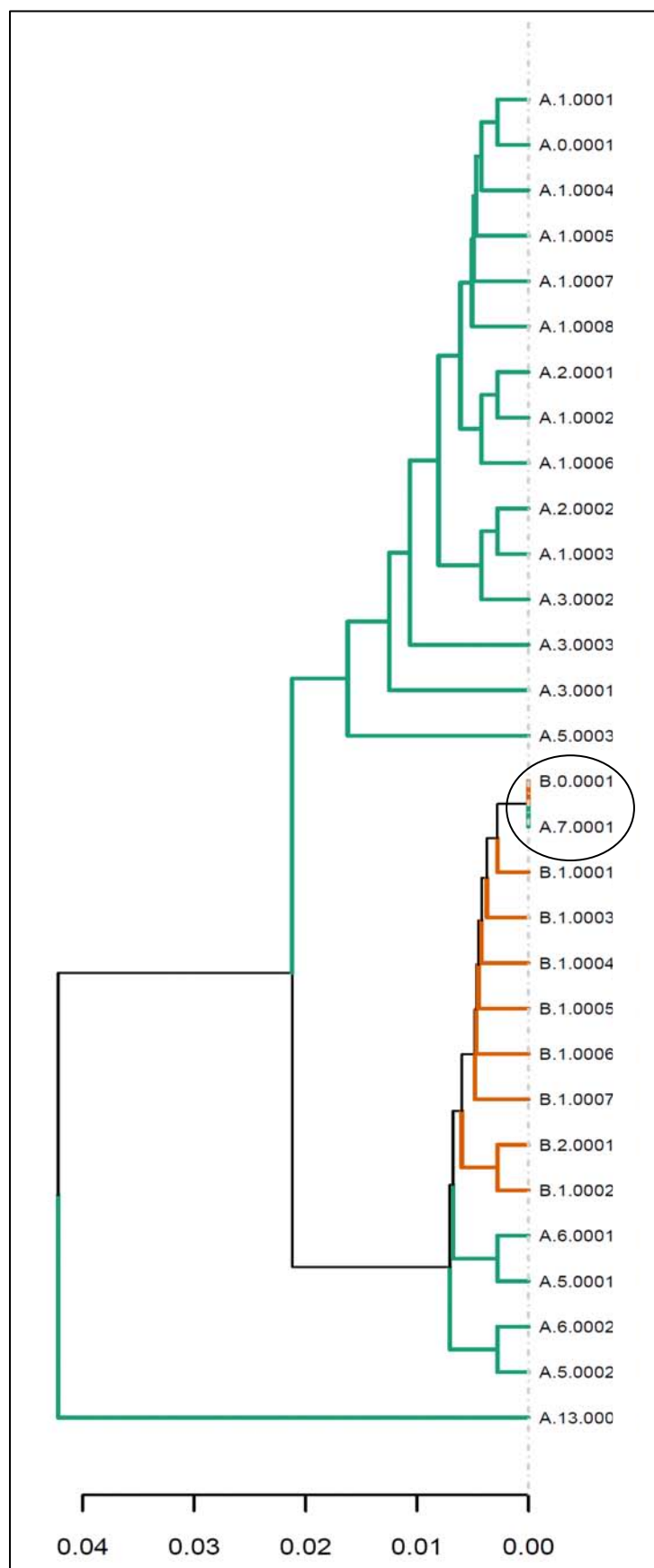
AMPLICÓN 3

AMPLICÓN 4

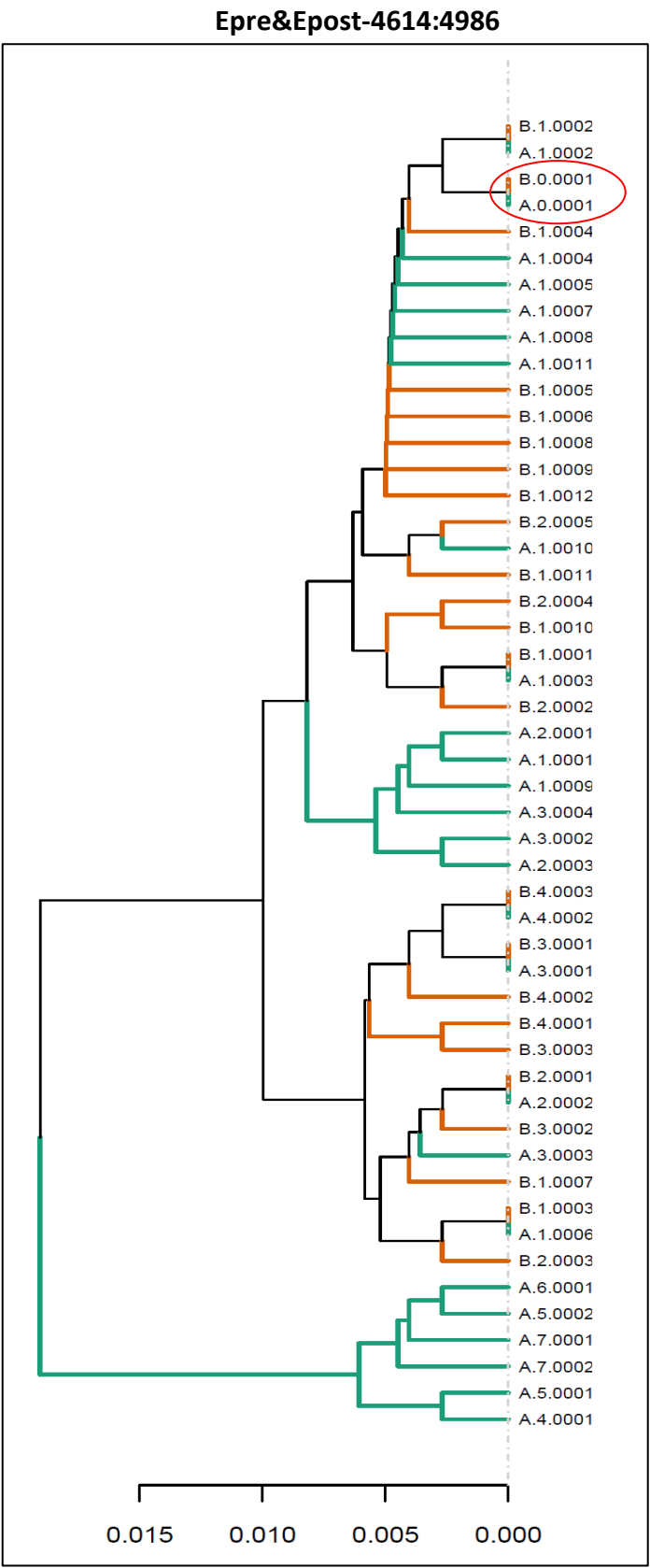


AMPLICÓN 5

Epre&Epost-4302:4661

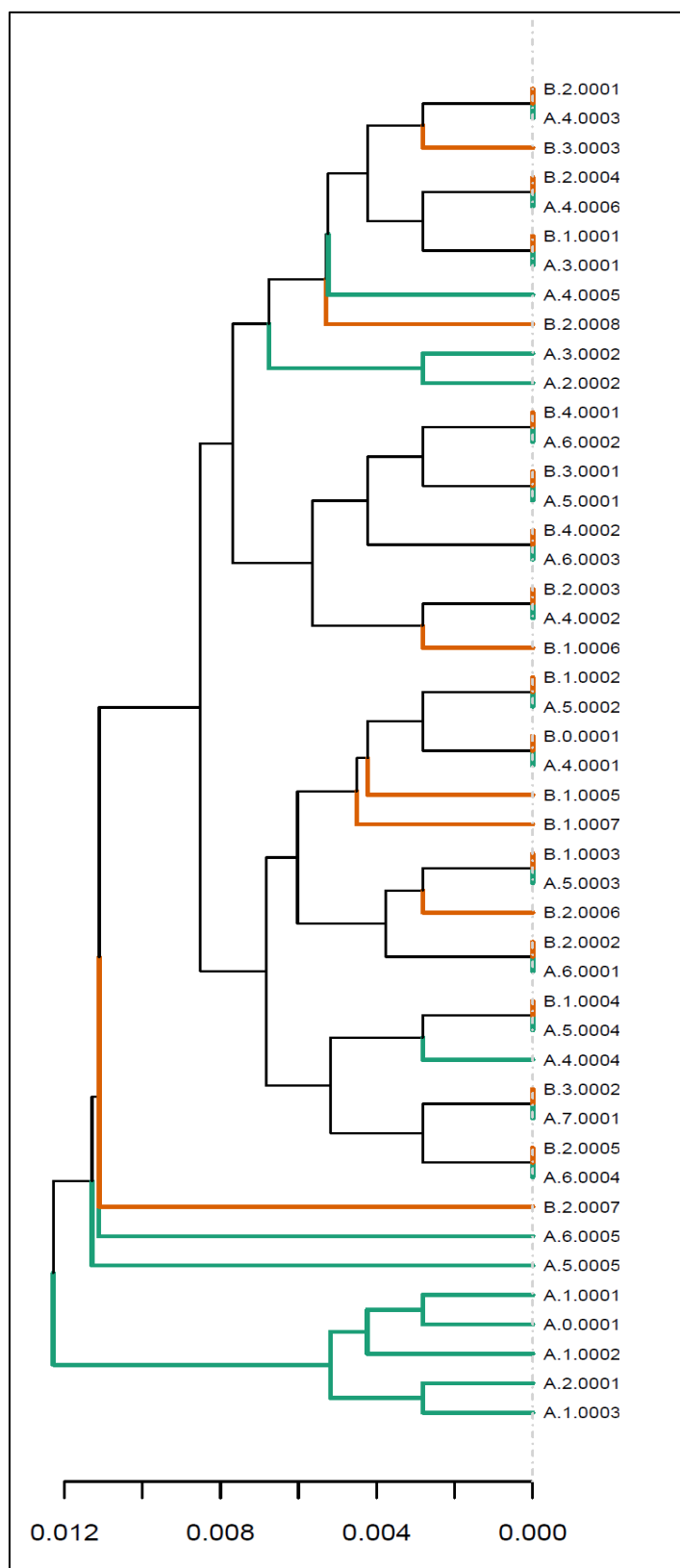


AMPLICÓN 6



AMPLICÓN 7

Epre&Epost-4926:5281



13 g

5.3.11 Comparación análisis filogenéticos y estructura de la densidad de mutación de los pacientes A y B

La observación de una estructura bimodal en los plots de densidad de mutación en el paciente A indicó un posible estado de evolución/competición dentro de la población del paciente crónico A en el aislado del año 1998. A partir de aislados secuenciales del paciente crónico obtenidos en 1993, 1996, 1997, 1998 y 2001, se comparó la secuencia consenso de estos aislados con los haplotipos de 1998 obtenidos en la secuenciación masiva (Figura 14).

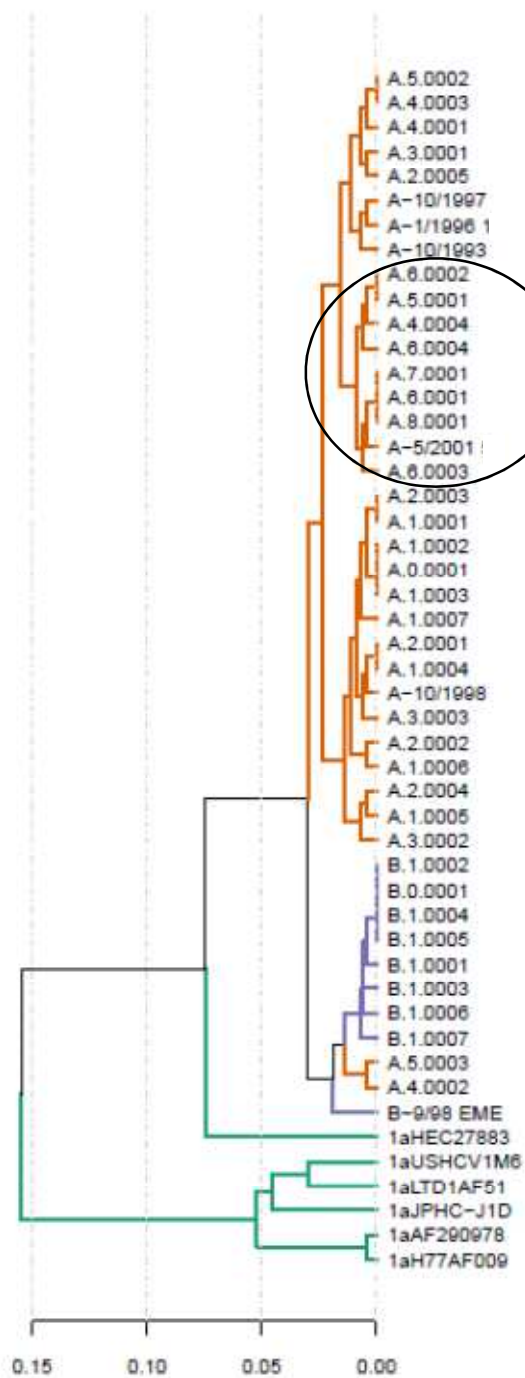
El estudio longitudinal de secuencias consenso de aislados de 1993, 96, 97, 98 y 2001, comparado con las secuencias obtenidas en 1998, mostró que los haplotipos de los amplicones 1, 2, 3, y 4 englobados en el segundo pico de densidad de distribución de la población viral (Figuras 14 a, b, c y d) se agruparon junto con las secuencia consenso del aislado del 2001. Por lo tanto, el segundo pico en los diagramas de densidad representa una subpoblación viral, que en 2001 se selecciona y que podría haber desplazado a la población viral mayoritaria observada en el año 1998. Esto indica que durante la fase crónica la población viral está evolucionando de manera continua y está sometido a cierta presión selectiva.

Figura 14. Comparación de secuencias de nucleótidos entre los pacientes A y B. Combinación de árboles filogenéticos UPGMA y Montserrat plots (páginas 96-102) . A la izquierda se representan los árboles filogenéticos y a la derecha los Montserrat plots en relación con el haplotipo dominante para los pacientes A (arriba) y B (abajo).

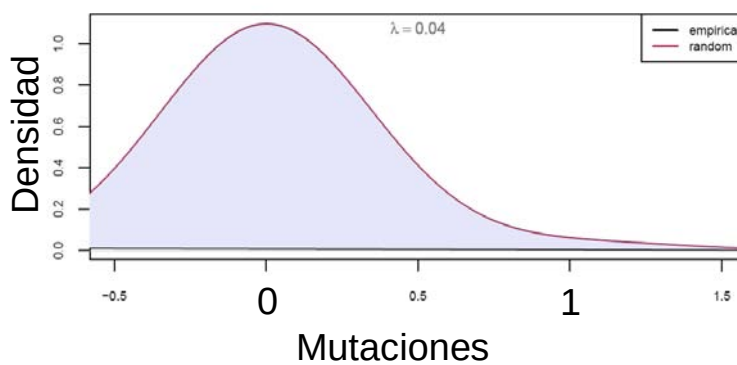
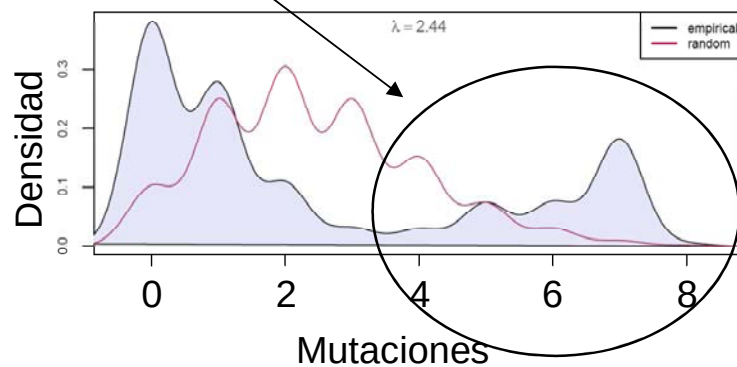
- El círculo continuo indica la agrupación de la secuencia consenso del paciente A del aislado del 2001 (A-5/2001) con secuencias de clones del paciente B aislados cerca del momento de transmisión en 1998.
- La doble flecha indica los mismos haplotipos en el árbol filogenético y en los Montserrat plots.

AMPLICÓN 1

UPGMA tree (K80; gamma=0.21): A&B-3420:3634

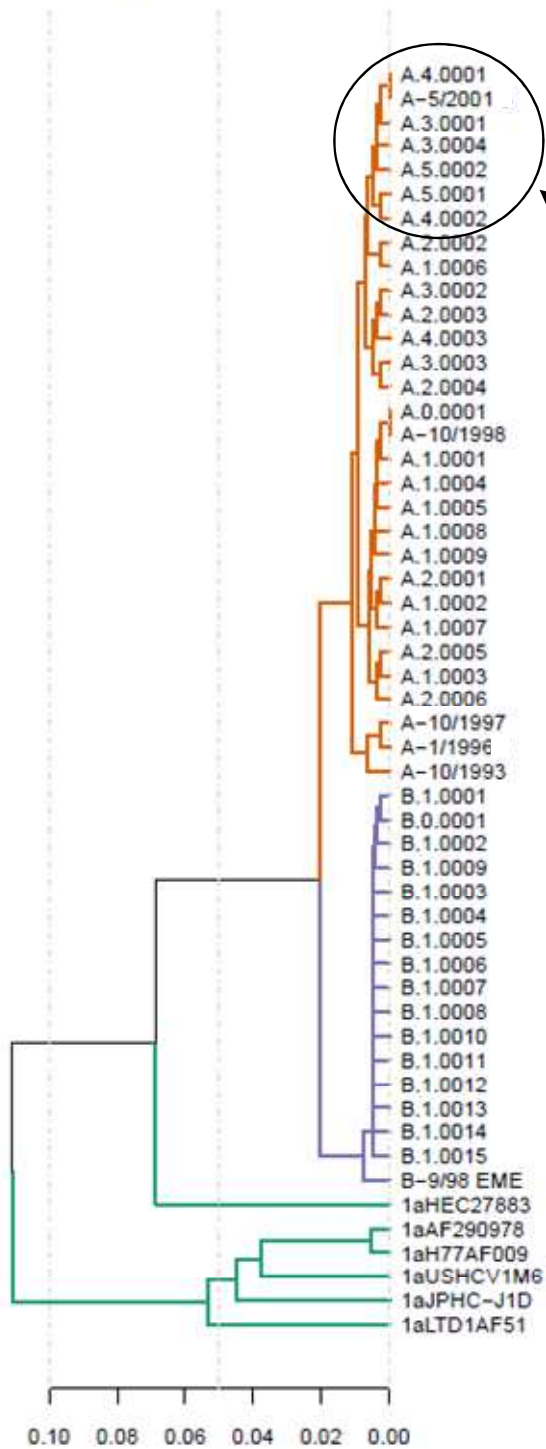


Haplotipos del paciente A (muestra de 1998) con 4-8 sustituciones que forman parte de la secuencia consenso 3 años después (en 2001)

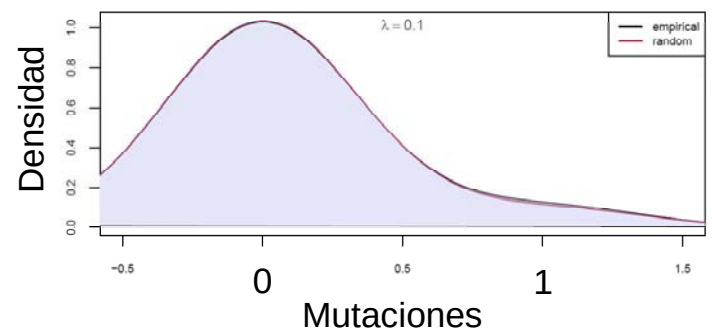
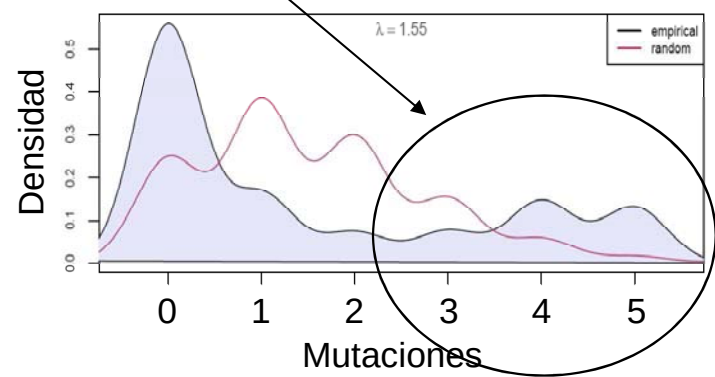


AMPLICÓN 2

UPGMA tree (K80; gamma=0.26): A&B-3555:3944

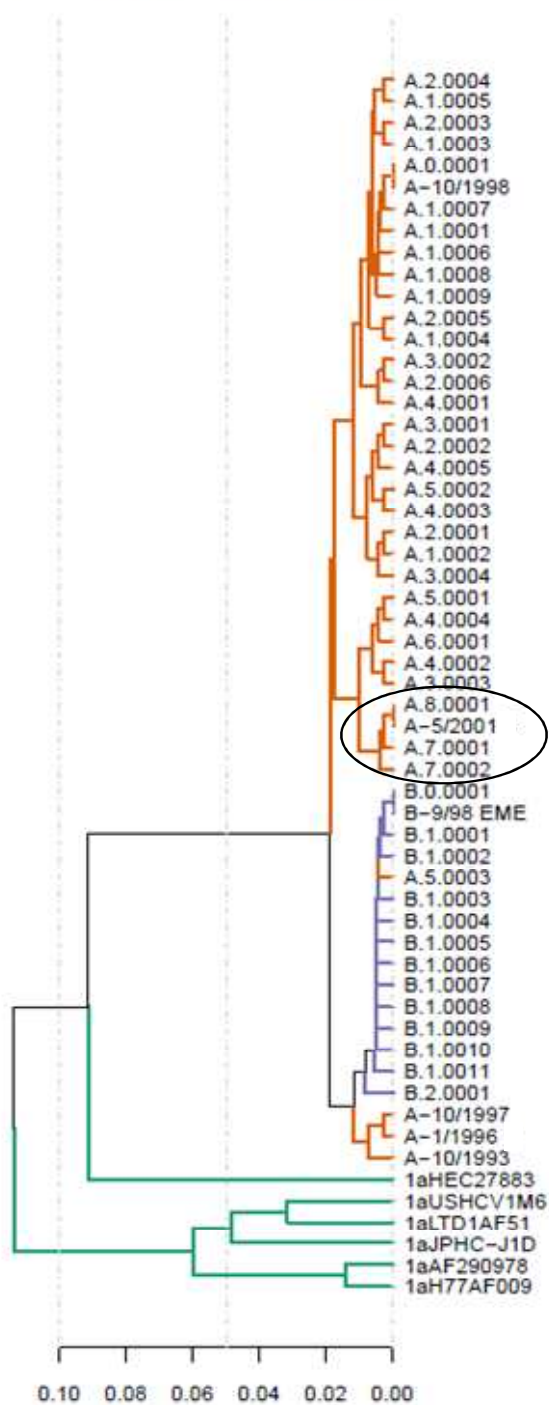


Haplotipos del paciente A (muestra de 1998) con 3-5 sustituciones que forman parte de la secuencia consenso 3 años después (en 2001)

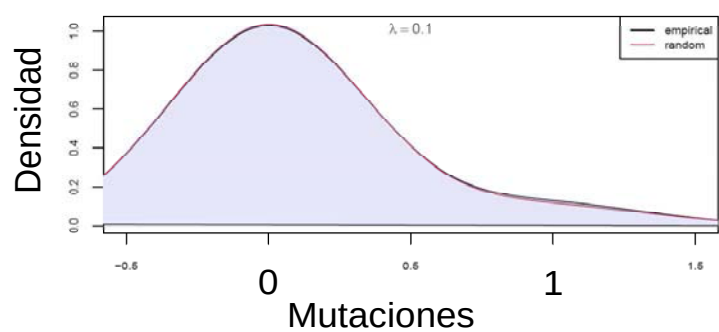
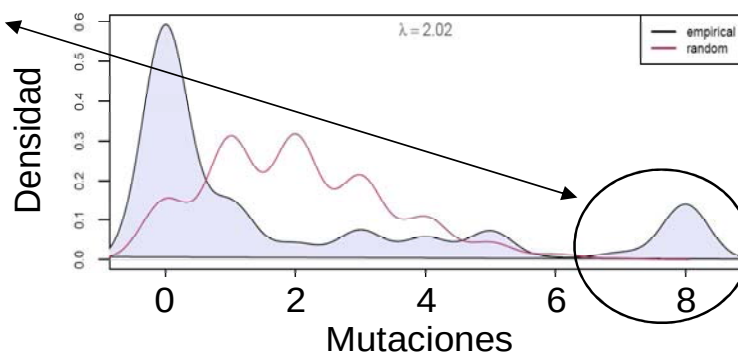


AMPLICÓN 3

UPGMA tree (T92): A&B-3840:4196

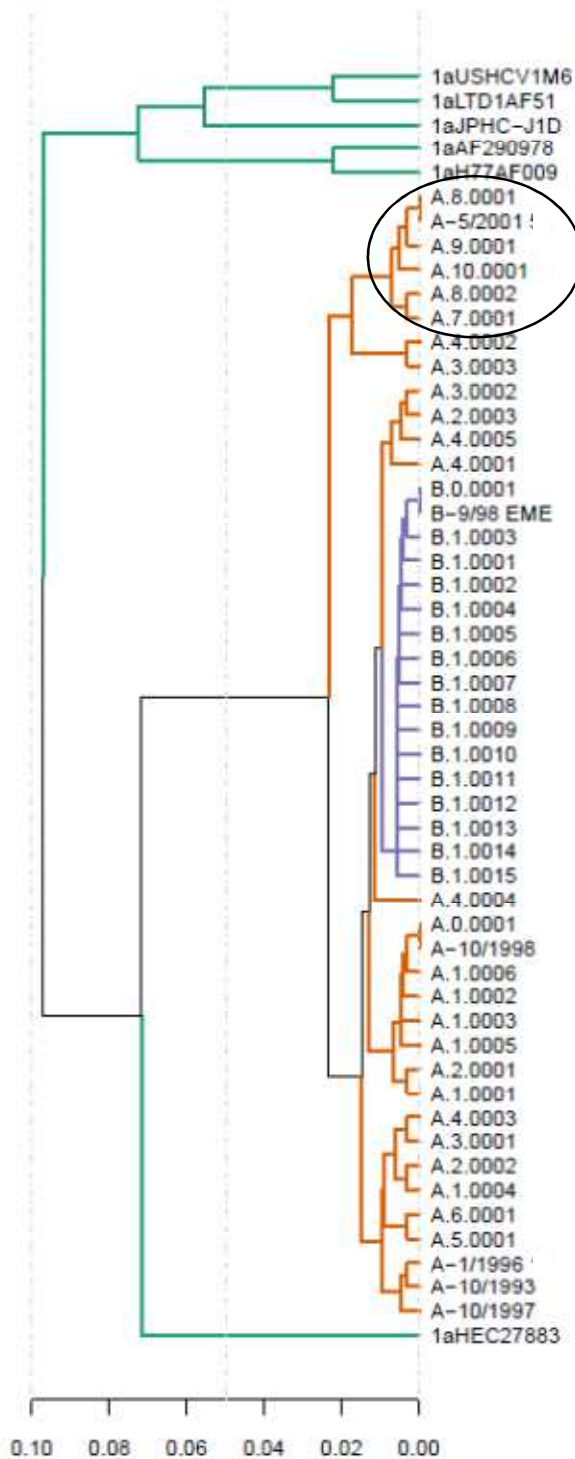


Haplotipos del paciente A (muestra de 1998) con 7-8 sustituciones que forman parte de la secuencia consenso 3 años después (en 2001)

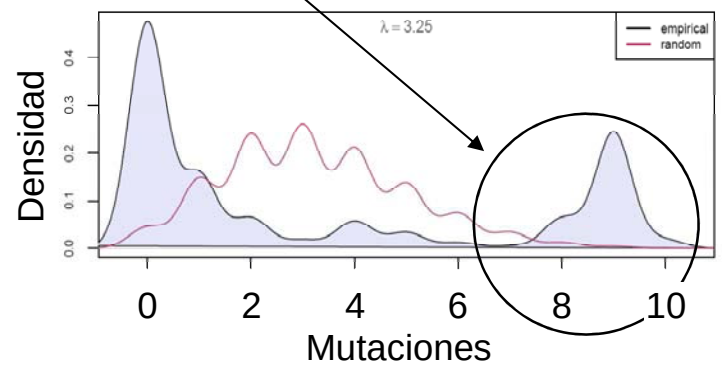


AMPLICÓN 5

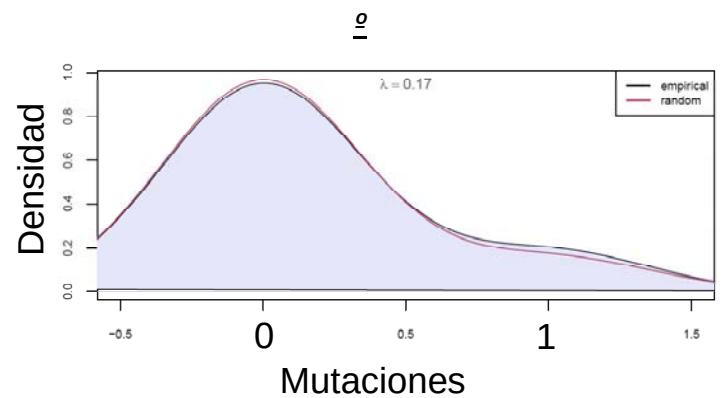
UPGMA tree (K80; gamma=0.22): A&B-4062:4391



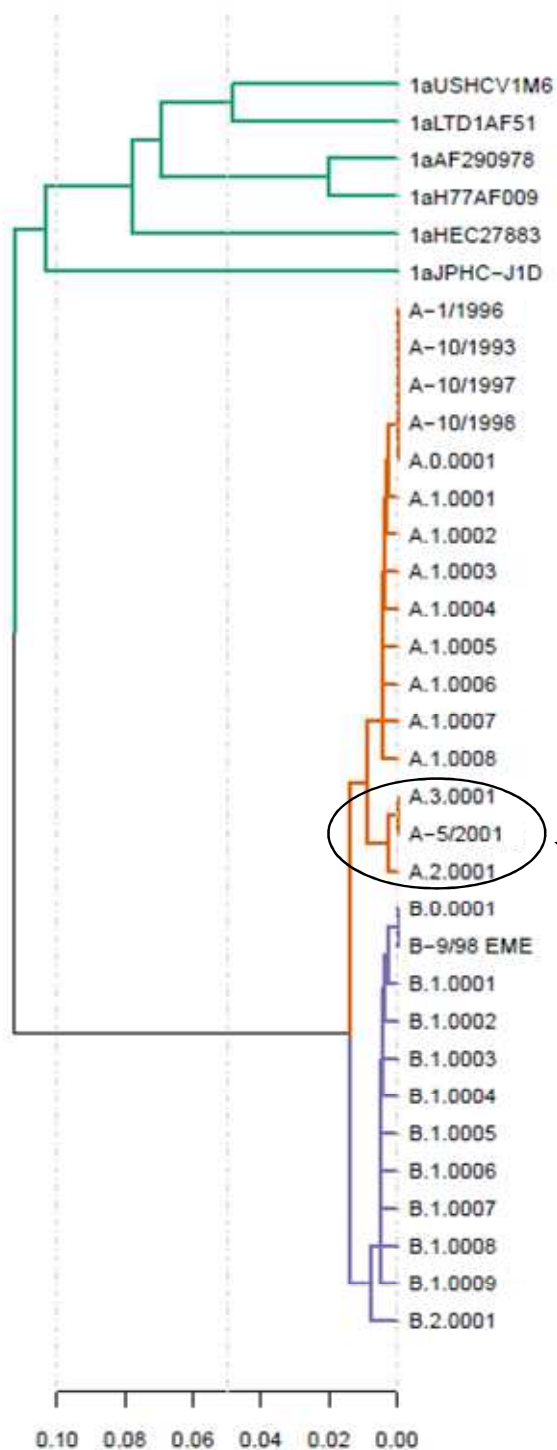
Haplotipos del paciente A (muestra de 1998) con 5-10 sustituciones que forman parte de la secuencia consenso 3 años después (en 2001)



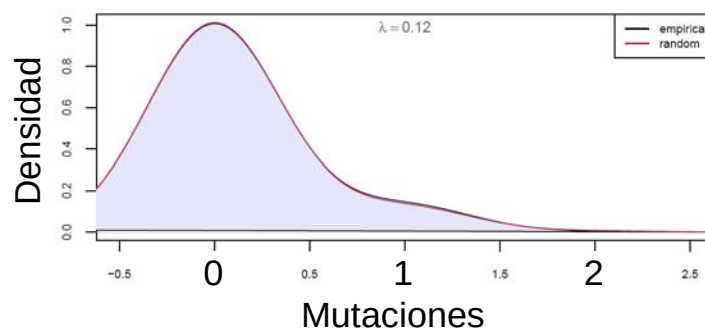
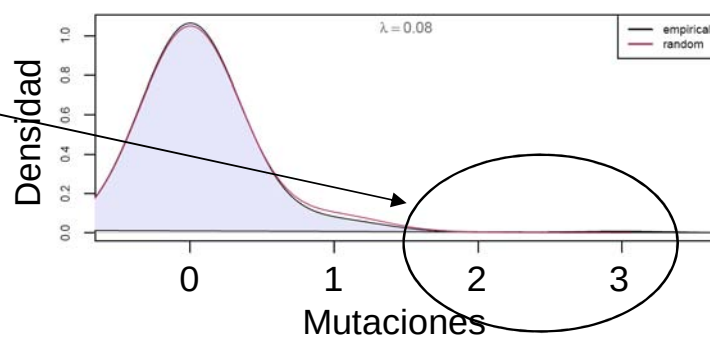
AMPLICÓN 5



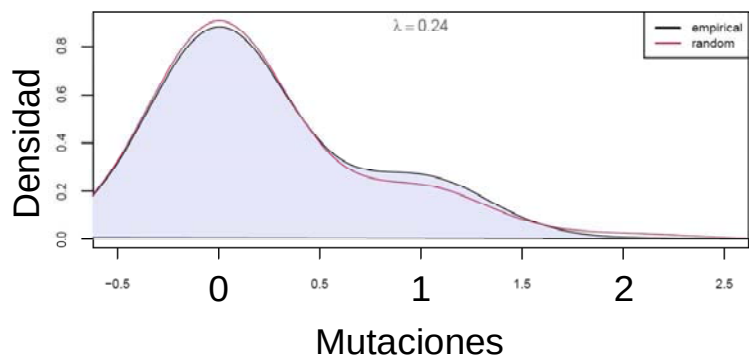
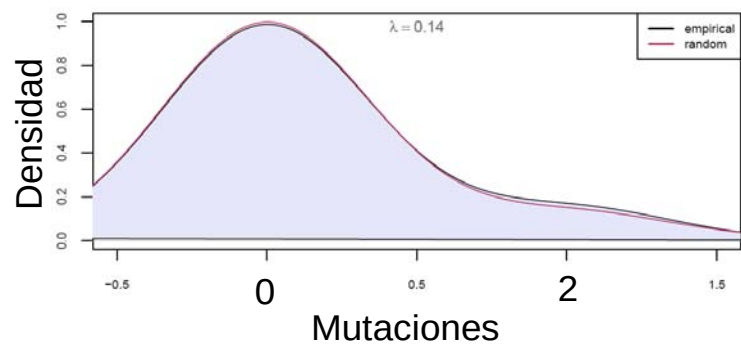
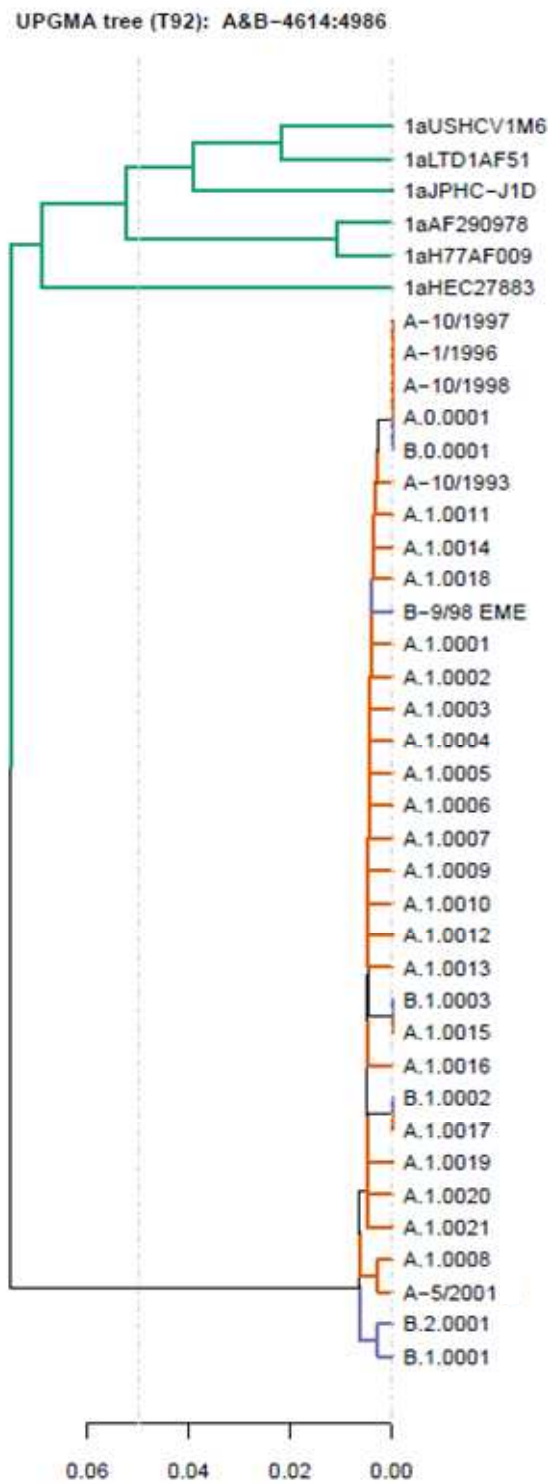
UPGMA tree (K80; gamma=0.21): A&B-4302:4661



Haplotipos del paciente A (muestra de 1998) con 2-3 sustituciones que forman parte de la secuencia consenso 3 años después (en 2001)

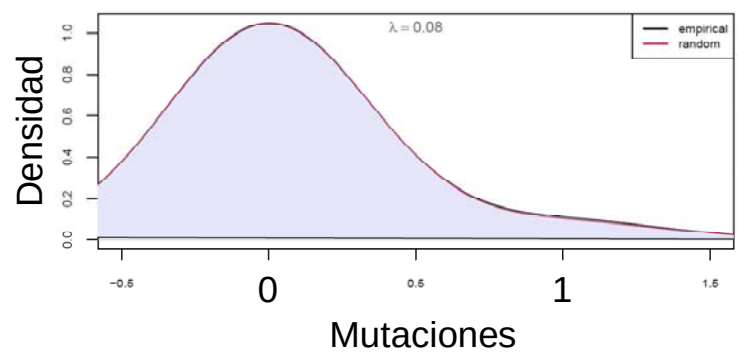
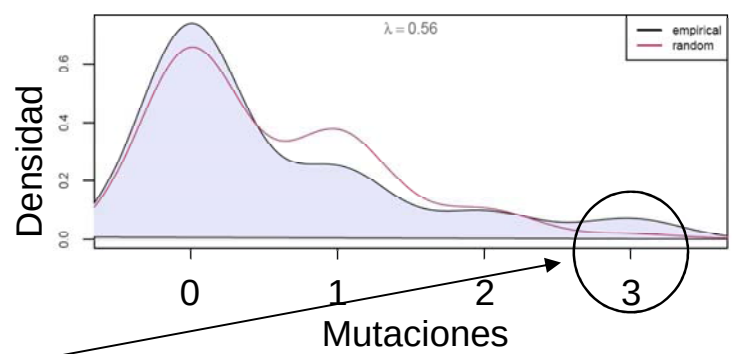
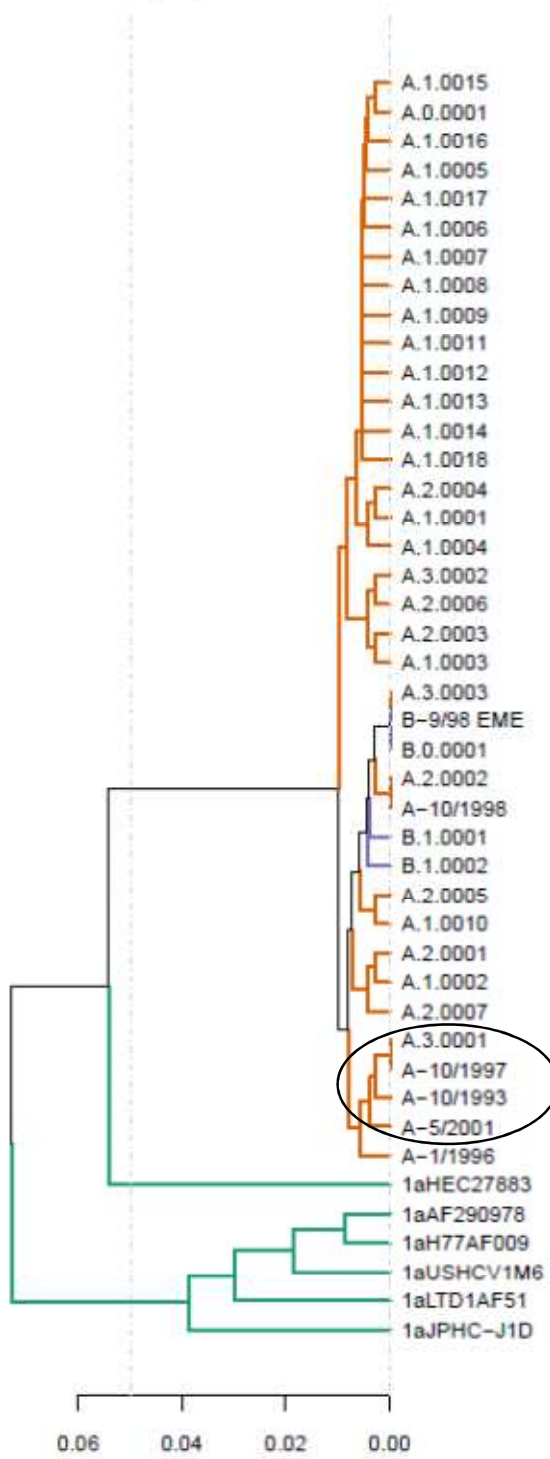


AMPLICÓN 6



AMPLICÓN 7

UPGMA tree (T92): A&B-4926:5281



5.3.12 Análisis de mutaciones de resistencia contra inhibidores de NS3 proteasa del VHC

El desarrollo de inhibidores de acción directa contra proteínas virales y en concreto contra NS3 del VHC ha generado interés en detectar la presencia de mutaciones que confieren resistencias y que pueden conducir al fallo en el tratamiento. Las posiciones aminoacídicas asociadas a resistencia a inhibidores de NS3 incluyen las posiciones: **C16** (111), **V36** (114-117), **A39**, **Q41** (118), **F43** (118), **T54** (114;115;117;119), **V55**, **R109** (71), **R155** (114;115;117;119), **A156** (71;115;117-119), **V158** (120), **G162** (71), **D168** (114), **V170** (114;118), **M175** (114) y las compensatorias **Q86** (71) y **P89** (71) (Tabla 12). Nuestros controles internos usando clones, han revelado que la secuenciación masiva permite detectar de manera fiable mutantes de resistencia presentes alrededor del 0,5% en la población, por debajo de este valor, habrán sustituciones que están en la población viral y sustituciones artefactuales producto del error, sobre todo de la polimerasa usada durante la PCR, difíciles de distinguir.

Para determinar si cualquiera de los 5 pacientes, no tratados previamente con inhibidores de NS3 proteasa, albergaban mutaciones preexistentes asociadas con resistencia a nuevos inhibidores de proteasa de VHC en los aislados de suero estudiados, 17 sitios en la región NS3 fueron examinados. El paciente C mostró una mutación de resistencia pre-existente en la región de NS3 (V55A), presente en la población viral con una frecuencia elevada de 3,12% y otra mutación con una frecuencia de 0.1% (A156T). Con respecto al aislado del paciente Epre también se analizó otra mutación con una frecuencia de 0.1% (D168E). En el resto de pacientes, se detectaron mutantes de resistencia en proporciones inferiores al 0.01% (Tabla 10). Hacen falta estudios más profundos para determinar cuál es el umbral a partir del cual una mutación detectada a nivel basal tiene un valor predictivo de fallo de respuesta al tratamiento antiviral con inhibidores específicos. Estudios preliminares sugieren que las mutaciones de resistencia en la proteasa presentes en menos del 1% de las quasiespecies virales en individuos naïve pueden emerger como

mutaciones dominantes sólo unos días después del tratamiento (121), así como el desarrollo de resistencia cruzada con compuestos relacionados.

Tabla 10. Mutaciones que confieren resistencia a inhibidores de NS3 proteasa del VHC

Posición mutaciones	Resistencia a compuestos inhibidores NS3-proteasa del VHC				
	Paciente A	Paciente B	Paciente C	Paciente D	Paciente Epre
C16S (1042)					C16S 0,009%
V36A/M/L/G (1062)					V36A 0,014%
A39V (1065)					
Q41R (1067)					Q41R 0,021%
F43S/I/V/C (1069)					F43I 0,0069%
T54S/A (1080)					T54A 0,018%; T54S 0,018 %
V55A (1081)			V55A 3,12%		
Q86R (1112)			Q86R 0,098%		
P89L (1115)			R109K 0,7%		P89L 0,081%
R109K (1135)					
S138T(1164)					
V158I (1184)					
R155K/T/I/S/M/Q/L/G (1181)	R155S 0,0092%	R155S 0,0168%	R155G 0,011%	R155Q 0,062%; R155G=0,041%	
A156N/T/S/I/G/F (1182)		A156V 0,042% 0,0168%	A156T 0,14%		A156T 0,035%
G162R (1188)					
D168A/V/E/G/N/T/Y/I/H (1194)			D168E 0,0043%		D168E 0,15%
V170A (1196)					
M175L (1201)					M175 0,035%
					A156T 0,042% G162R 0,025% D168G 0,017%; D168Y 0,0085%
					P89L 0,098%

DISCUSIÓN

6 DISCUSIÓN

La transmisión de virus de un paciente a otro es una oportunidad única para estudiar factores virales y del huésped que pueden influir en el curso natural de la infección por VHC hacia resolución espontánea o cronicidad.

En el presente trabajo se han estudiado 3 casos de transmisión del VHC, que a priori, podrían involucrar un tamaño de inóculo viral diferente. Un caso de transmisión por cuello de botella genético demostrado (estudio 1), un caso de transmisión nosocomial por TAC con contraste en que se desconoce el tamaño de inóculo transferido y un tercer caso de contacto masivo después de un trasplante de hígado en un paciente crónico VHC+.

En los 3 casos, no existieron diferencias significativas en el título viral comparando cada miembro de la pareja: 10^6 UI/ml. Este dato permitió comparar la variabilidad en cada pareja sin riesgo de sesgo debido a la carga viral.

El objetivo de este trabajo fue comprobar cómo se comportaban las poblaciones virales después de la transmisión, comparados con su fuente y con el resultado de la infección, teniendo en cuenta datos genéticos del paciente, datos genómicos del virus, así como la respuesta inmune.

6.1 Factores del paciente (IL28 y HLA)

Respecto a los datos genéticos de IL28 y HLA de los pacientes, se observó que el paciente agudo curado B y el paciente agudo D, que estaba en fase de iniciar persistencia, presentaron el genotipo favorable CC. Por tanto, presentar un genotipo CC que se ha asociado a curación espontánea o resolución de la infección por VHC, no es un requisito suficiente para determinar el resultado de la evolución de la infección por VHC. En relación a los alelos de HLA clase II que se relacionan con respuesta inmune anti-VHC específica, tanto A (crónico) como B (curado) presentaron haplotipos asociados a curación. Por tanto, presentar un haplotipo favorable no permite predecir si el paciente se curará espontáneamente o no. En

conclusión, el SNP de IL28B y el haplotipo de HLA por si sólo son datos genómicos insuficientes para predecir la resolución espontánea de la infección por VHC en estos casos concretos.

6.2 Factores virales

- Nucleótidos

El inicio de la infección en un individuo naïve o que recibe una inmunosupresión, suponen situaciones interesantes de estudio donde se predice que los virus con mayor fitness pueden proliferar con mayor eficacia que los competidores, hecho que se refleja en una homogeneización de la población viral, hecho que ya ha sido descrito en la literatura.

En nuestro estudio, la diferencia en el grado de homogeneidad fue clara entre las 3 situaciones clínicas y sobre todo entre pacientes agudos y crónicos, estos últimos presentando una gran heterogeneidad. Respecto a las diferentes situaciones clínicas, después de un cuello de botella (estudio 1) y en plena fase aguda (caso A/B), la homogeneización fue mucho más acentuada (valor de diversidad de nucleótidos A/B=9,64) que en el caso del trasplante (caso Epre/Epost=3,03) o de la transmisión por TAC con contraste (caso C/D=2,13). En los 3 casos estudiados, tan sólo el paciente B que se curó espontáneamente, es el que presentó la población viral más homogénea.

En el 2º caso (caso C/D), los datos de complejidad junto con los datos de filogenia molecular, indicaron una situación de transmisión de un número importante de partículas (ya que en fase aguda el paciente D presentó una población compleja similar a su partner paciente C), por tanto, la transmisión del VHC por TAC con contraste puede suponer un riesgo importante de transmisión del VHC en cantidades suficientes para causar cronicidad en un porcentaje de pacientes alto, más similar a una transfusión de sangre que a un pinchazo accidental con aguja infectada.

La gran homogeneización observada en el paciente curado B puede haber facilitado la eliminación del virus. Se ha descrito en la literatura que una respuesta inmune efectiva dirigida hacia una población homogénea tiene más probabilidades de éxito que frente a una población heterogénea de virus, donde es más probable que aparezcan mutantes de escape. Farci et al. demostraron que pacientes infectados por el virus que presentaban escasa evolución durante la etapa aguda de la infección se recuperaron completamente, mientras que los pacientes en los que el virus evolucionó rápidamente desarrollaron infección crónica (55). Varios estudios han asociado una complejidad alta del espectro de mutantes con la ausencia de la resolución de la infección por VHC, ya sea espontánea o mediada por tratamiento (122;123). En el caso del trasplante de hígado por VHC, numerosos estudios muestran que la reinfección y selección de variantes virales son procesos muy dinámicos después de la reperusión y la población viral del VHC se vuelve más homogénea, incluso después de cambios importantes en su composición, pero en este caso la cronificación puede ser favorecida por la inmunosupresión a la que están sometidos estos pacientes (124-126).

- [Aminoácidos](#)

Sorprendentemente, la diferencia en complejidad de las quasiespecies a nivel de nucleótidos entre pacientes agudos y crónicos, no se reflejó a nivel de aminoácidos. En los 3 pacientes crónicos, la gran proporción de mutaciones nucleotídicas observadas en pacientes crónicos no se vio reflejada a nivel de aminoácidos. El valor más espectacular se observó en la pareja A/B donde el paciente agudo B presentó incluso mayor número de mutaciones en aminoácidos que en la fuente de infección (paciente A). De los 7 amplicones estudiados por secuenciación masiva, el número 7 (1529aa-1646aa), situado en el fragmento carboxi-terminal de la NS3, fue el que presentó una frecuencia de cambios no sinónimos más alta en el paciente crónico A y la más baja en el paciente curado B, todo lo contrario de lo observado en el resto de amplicones (del nº 1 al nº 6), en los que la mayoría de las mutaciones fueron sinónimas. Por tanto, los datos de

secuenciación masiva nos dirigen hacia la región terminal de la helicasa como una región clave de comportamiento diferencial entre los pacientes A y B. El estudio de respuesta inmune confirmó el interés de esta región viral, puesto que en los pacientes A y B, la técnica de Elispot demostró que el paciente crónico tenía una respuesta inmune CD4 específica indetectable en esta región carboxi-terminal de NS3, mientras que por el contrario, la paciente curada fue precisamente en esta región, donde presentó una mayor respuesta específica tipo CD4. Estos resultados sugieren que esta zona carboxi-terminal de NS3 puede ser relevante para explicar el resultado de la infección (curación o cronicidad). Cabe destacar que a nivel de mutación, esta región fue precisamente, donde la población viral del paciente crónico A aceptó una mayor proporción de mutaciones no sinónimas, sugiriendo que a pesar de la baja respuesta inmune, estos cambios podrían conferir una ventaja al virus para mantener la cronicidad.

En el paciente agudo B, la fuerte respuesta inmune fue acompañada de un mínimo o nulo cambio de aminoácidos. Esta situación podría parecer paradójica, en el sentido de que siempre se ha hablado de que una mayor presión selectiva induce al virus a mutar, pero existe una explicación que fue sugerida por Farci et al. que propuso que esta gran homogeneización de la población puede ser un reflejo de la respuesta inmune efectiva, que supone un paso previo a la eliminación de la infección viral (55). Esta observación fue similar a previas publicaciones que han documentado que una respuesta persistente y multi-específica dirigida a epítomos de NS3 de células T fue más eficaz en los pacientes que resolvieron la infección (87;89;93;127;128). En contraste, la cronicidad a menudo se caracteriza por la anergia de células T causadas por la inactivación de la respuesta de las células T CD4+ específicas de VHC que impide la eliminación viral (129;130). Estudios de modelos en ratones infectados con otros virus han demostrado que la inactivación funcional de las células T CD4+ puede ocurrir (a nivel celular e independientemente de las células T reguladoras o de células presentadoras de antígeno-CPA-) al principio de su transición a la persistencia viral, que llevan a su consiguiente incapacidad para mantener la función de las células T citotóxicas, facilitando así la persistencia, aunque un reservorio pequeño de las células T CD4+ activas pueden

persistir durante la cronicidad (131). Recientemente se ha demostrado, que una vez, que la cronicidad se ha establecido, la respuesta específica CD4 inactiva, se puede restaurar, mediante la expansión ex vivo de células CD154+ específicas de NS3 de pacientes crónicos, la producción de IFN- γ e IL-2 y la proliferación a niveles similares a las de pacientes que resolvieron la infección espontáneamente, cultivando las células CD4+ en ausencia de antígeno específico (132). La acumulación de cambios de aminoácidos en epítomos de células T de la región carboxi-terminal de NS3, junto a la ausencia detectable de CD4+ activadas, sugiere un estadio de evolución viral continuada para mantener el estado de tolerancia inmunológica como consecuencia de la competencia entre la entrada mutacional del VHC y la respuesta del sistema inmune del individuo infectado, más que aparición de mutaciones de escape a la respuesta inmune (133;134). Al menos cuando una infección está bien establecida, el VHC puede mostrar una dinámica de escape frente a anticuerpos neutralizantes y células T citotóxicas, como uno de los mecanismos evolutivos que contribuyen a la persistencia (68;135;136).

En conclusión, la región carboxi terminal de NS3 de VHC presentó un comportamiento diferente cuando se comparó el paciente crónico A y su pareja curada B, revelado por análisis de secuenciación masiva y evidenciado a nivel fenotípico por una respuesta fuerte multiespecífica CD4 contra epítomos de células T situada en la región pertinente.

- Density plots (Montserrat plots)

Durante la caracterización de las poblaciones virales, se construyeron plots de densidad que reflejan si las poblacionales virales están alejadas o cercanas a lo esperado si las mutaciones fuesen fijadas al azar. En los 3 pacientes infectados “de novo” los plots de densidad presentaron una distribución muy cercana al azar. Por el contrario, el paciente crónico A, con más de 8 años de cronicidad, apareció una distribución bimodal alejada del azar. Esta distribución fue aún más extrema en el paciente Epre cirrótico donde, en algún amplicón apareció una distribución

trimodal con subpoblaciones con más de 15 mutaciones respecto a la más representada (máster). Esta gran complejidad puede ser un reflejo del daño hepático, con islas de hepatocitos infectadas aisladas por una estructura fibrosa del hígado. El paciente C fue un caso particular, pues a pesar de ser crónico, presentó una distribución de azar. Sería interesante estudiar si la distribución en plots de densidad se relaciona con daño hepático.

- *Pareja de pacientes A/B*

En el caso del paciente A, en el que la distribución de la población viral se alejó del azar, se observó que la mayoría de las mutaciones fueron sinónimas, sugiriendo que tenía una quasiespecies muy bien adaptada, permitiendo pocos cambios drásticos. Todos los cambios de nt son cambios genotípicos pero aquellos que inducen cambio de aminoácido, pueden afectar a la actividad de la proteína y por tanto, pueden implicar cambios fenotípicos. En términos de biología de sistemas, se define a una población como robusta si su función o estructura persiste, a pesar de las mutaciones, mientras que se define como evolucionable si las mutaciones pueden producir cambios fenotípicos (137). En nuestro caso podemos definir robustez como la constancia en la secuencia de aminoácidos a pesar de la mutación genotípica. En el caso del paciente A, si nos restringimos al nivel fenotípico (cambios de aminoácidos), la población viral en el aislado de 1998, parecía que estaba en un estado de estasis evolutiva, donde los cambios de aminoácidos fueron seleccionados negativamente, lo que en evolución se conoce como selección purificadora. No obstante, a nivel genómico (cambio de nucleótidos) presentó una frecuencia de mutación elevada, y por tanto, un potencial de evolución alto. El hecho de poseer muestras seriadas de suero del paciente A, permitió comparar la población viral de 1998 con la secuencia consenso de 2001. Sorprendentemente, la secuencia consenso de 2001 derivó de una subpoblación detectada en 1998 y que tenía más de 3 mutaciones (entre 3 y 10) con respecto a la secuencia mayoritaria encontrada en el aislado de 1998, indicando que la población de 1998 estaba en continua evolución. En los plots de

densidad (Montserrat plots) esta población conformó un 2º pico. Los mutantes de este pico se agruparon junto al aislado de 2001 en los árboles filogenéticos.

Por tanto, estos resultados sugieren que bajo una población altamente robusta a nivel fenotípico, pueden acumularse cambios a nivel genotípico que permiten a la población, en un momento dado, evolucionar. Si esto se extrapola a términos de “fitness landscapes” (paisajes de fitness”), se podría hablar de que en 1998, la población viral del paciente A estaba situada alrededor de un pico de fitness, bien adaptada, pero que debido a la continua acumulación de mutaciones, esto pudo haber permitido a la población viral expandir sus límites, facilitando que encontrase un pico de fitness mejor, sobre todo, en un ambiente cambiante, conduciendo finalmente, a un cambio evolutivo, como el observado en el año 2001.

- *Pacientes C/D y Epre/Epost*

En los otros 2 casos de transmisión del VHC, en los casos D y Epost de nueva infección, la población viral evolucionó hacia cronicidad.

A destacar que en el aislado post-trasplante, en una situación de inmunosupresión, la población viral a los 3 meses, presentó una estructura más cercana a una fase de cronicidad avanzada que a una fase aguda previa a la cronicidad. Este hecho podría ser debido a que en un medio donde la presión inmune está limitada, no existe presión selectiva, el virus puede evolucionar y adaptarse de una manera más rápida al nuevo hígado.

- *Análisis filogenéticos*

El modo de transmisión del VHC asociado a la cantidad de virus transmitidos podría ser uno de los factores clave para la transición de infección aguda a persistente. La entrada masiva de virus se ha asociado a un mayor riesgo de cronicidad, mientras que la entrada de un inóculo pequeño (cuello de botella genético), a una menor capacidad del virus para persistir.

Nuestros datos confirman estas observaciones, ya que en la transmisión por cuello de botella (A/B) la paciente aguda B se curó, mientras que en los casos C/D y Epre/Epost se estableció cronicidad. Es interesante destacar que la pareja de aislados C y D, fue la que presentó mayor similitud entre poblaciones, sugiriendo que la infección del paciente D pudo iniciarse a partir de una parte representativa de la población viral del paciente crónico C, y ello facilitar el desarrollo de una infección crónica. En cambio, las poblaciones virales de los pacientes A y B, caso en el que se produjo una transmisión por cuello de botella, estaban más separadas filogenéticamente, mientras que el caso del trasplante, la infección masiva del nuevo hígado se caracterizó por una estructura poblacional más cercana a un estado de cronicidad avanzada que en una fase aguda en un paciente inmunocompetente. Esta estructura poblacional podría ser el resultado de la selección de las variantes virales más competitivas (con mayor fitness), acompañado por el estado de inmunosupresión del paciente. Es decir, tres casos, tres situaciones clínicas diferentes, pero se necesitaría corroborar estos estudios con otras parejas similares. Esto es difícil en los casos A/B, C/D, pero puede ser interesante realizarlo en pacientes transplantados.

- *Estudio de mutaciones de resistencia*

En el caso de los 6 pacientes del estudio, ninguno de los cuales recibió tratamiento, se detectó en suero en el paciente C una mutación de resistencia (V55A) con una frecuencia de 3,12% y otra A156T a una frecuencia de 0,1% (A156T). En el aislado del paciente Epre también se detectó la mutación D168E con una frecuencia de 0,1%. En el paciente A la mutación A156T (estudio 2-página 36), que estaba fluctuando en un 0.78% en la población viral en un aislado de biopsia hepática, en el suero correspondiente al año 98 no se detectó en valores de frecuencia >0,1%. Es importante realizar más estudios para determinar cuál es la frecuencia en la que están presentes estas mutaciones en la población viral, si se mantienen en el tiempo, y cuál es la frecuencia mínima clínicamente relevante cuando se pretende tratar al paciente con los inhibidores específicos de NS3.

En la actualidad el procedimiento de secuenciación masiva es una herramienta que permite cuantificar la complejidad de las poblaciones virales y estudiar la composición del espectro de mutantes en profundidad. Aunque se conocen numerosas interacciones entre factores virales y celulares (esenciales para la replicación y/o en la respuesta anti-viral intracelular y su evasión), aún no ha sido posible establecer un mapa poblacional de las variantes virales de cada paciente en función de polimorfismos genéticos virales y del huésped, que permitan predecir la respuesta terapéutica con suficiente precisión para individualizar el régimen terapéutico. Por tanto, el conocimiento detallado de factores virales y del paciente, proporcionará una herramienta útil para la comprensión de la infección por VHC y proveer perspectivas para el diagnóstico y diseño de una terapia personalizada.

CONCLUSIONES

7 CONCLUSIONES

- 1- La transmisión del VHC por vía sexual supone un cuello de botella genético.
- 2- La transmisión por cuello de botella supone una homogeneización muy importante de la población viral.
- 3- Pacientes antes del tratamiento pueden presentar mutaciones de resistencia que dificulten el tratamiento con inhibidores específicos.
- 4- El inicio de la infección supone una homogeneización de la población viral que es más acentuada cuanto menor es el inóculo viral infeccioso.
- 5- En la infección crónica predominan los cambios sinónimos.
- 6- La mutación se distribuye de manera desigual a lo largo del gen NS3.
- 7- La región carboxi-terminal de NS3 puede jugar un papel clave en la resolución/cronicidad de la infección por VHC.
- 8- La R.I. CD4+ frente a diferentes epítomos de la región carboxi-terminal acompañada de una homogeneización a nivel de aas se asocia a resolución de la infección.
- 9- La falta de R.I. CD4+ y cambio de aa en la región carboxi-terminal se detecta en la situación de infección crónica por VHC en un paciente.
- 10- La población viral en fase crónica, a pesar de presentar pocos cambios de aa, mantiene un nivel de variación a nivel de nt que indica una evolución viral basal continuada en el tiempo.

- 11- La presencia de mutaciones de resistencia con baja frecuencia ($<1\%$) en la población viral, antes de un tratamiento, podría ser un factor clave para definir la terapia antiviral personalizada más efectiva. Para ello es importante establecer cuál es la frecuencia mínima a la que se detecta un mutante que tiene relevancia clínica mediante la técnica de UDPS.

**MATERIALES Y
MÉTODOS**

8 MATERIALES Y MÉTODOS

8.1 Pacientes

Paciente A: paciente con historial de drogas intravenoso encontrado anti-VHC positivo en 1993 durante una donación voluntaria y al cual se le hizo un seguimiento en el tiempo. Este paciente tenía infección persistente con un genotipo 1a, que se confirmó por secuenciación de las regiones 5' UTR, E2PePHD y NS3 y una carga viral en suero de 10^6 IU/ml. Presentó niveles elevados de alanina aminotransferasa (ALT) en el tiempo con una media de 178 UI/ml y mediante biopsia hepática se demostró que padecía una hepatitis crónica activa moderada. Este paciente recibió un tratamiento con interferón estándar y ribavirina (IFN+RBV) durante 12 meses (entre Febrero 1996 y Marzo de 1997) con bioquímica transitoria pero no respuesta virológica. En Noviembre de 2006 fue sometido a una biopsia hepática de control.

Paciente B: pareja del paciente A, desarrolló en agosto de 1998 una hepatitis aguda sintomática. El desarrollo de estudios filogenéticos demostró que el paciente A transmitió la infección por vía sexual a su pareja la paciente B aproximadamente en la fecha de Junio de 1998. El episodio ocurrió cerca de siete semanas después de tener relaciones sexuales sin protección con el paciente A, poco después de la eliminación de un dispositivo intrauterino, debido a hemorragia vaginal. En el paciente B los niveles de ALT se normalizaron en 6 semanas y el ARN del VHC se eliminó 8 semanas después del inicio de los síntomas. Las muestras utilizadas para demostrar la transmisión, fue en septiembre de 1998, después de superar el pico de ALT en el momento de la infección aguda en el paciente B y en octubre de 1998 en el paciente A.

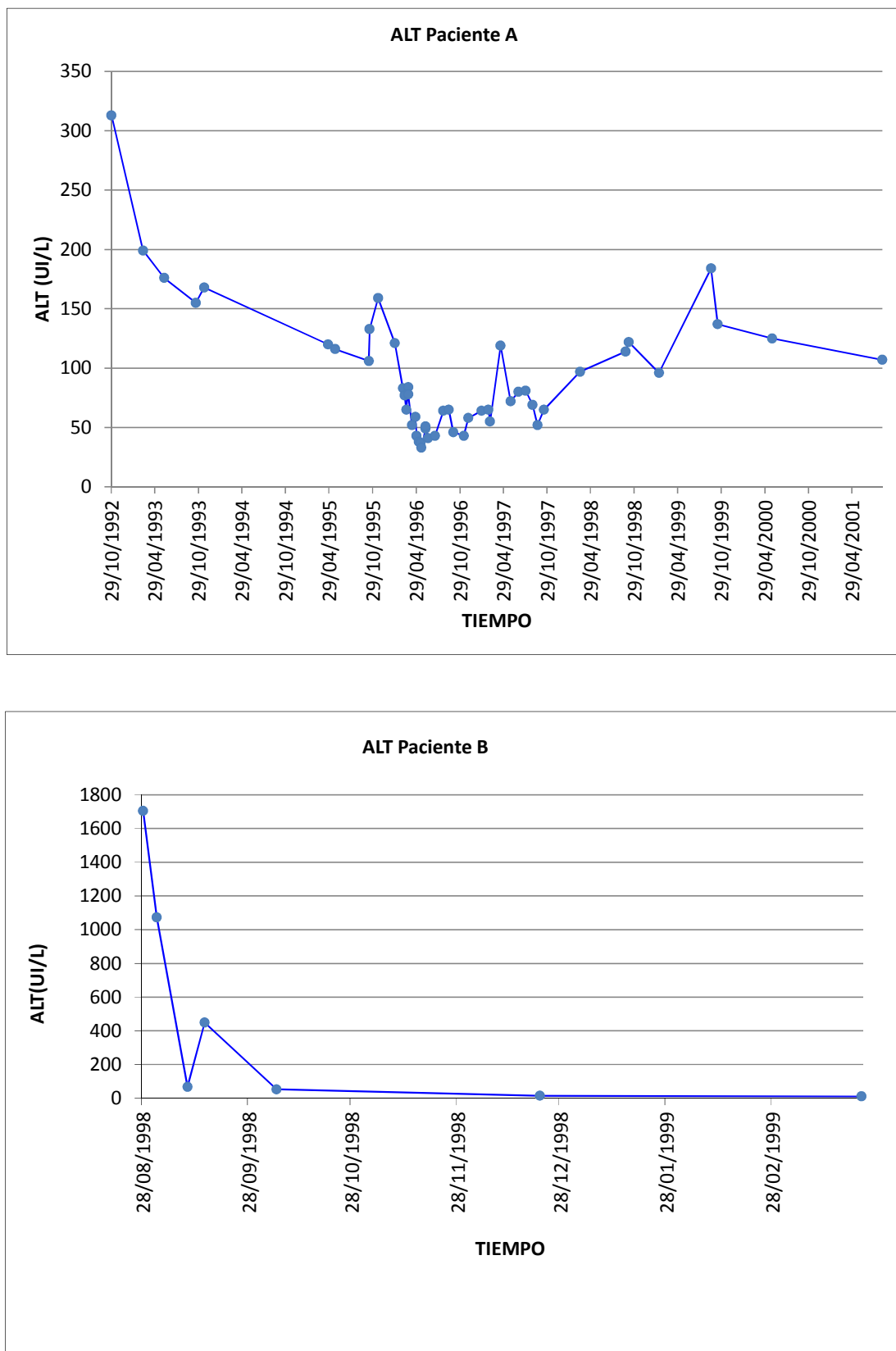


Figura 16. Valores de alaninoaminotransferasas (ALT) en pacientes A (gráfico superior) y B (gráfico inferior).

Paciente C: paciente con enfermedad de Crohn ileocólica y con infección por virus C genotipo 1b que en Junio de 2004 fue sometido a una tomografía axial computerizada (T.A.C.) por contraste justo antes que el paciente D. Este paciente falleció años antes de que se le realizaran pruebas de IL28 o HLA. Dicho paciente fue objeto de una exploración abdominal, la TC con contraste antes de paciente D.

Paciente D: paciente varón de 29 años de edad, con enfermedad activa íleo-cólica de Crohn. Se sometió a un TAC con contraste en 2004. Las serologías de hepatitis B y C durante el ingreso fueron negativas. Siete semanas más tarde, fue admitido en el hospital debido a náuseas, fatiga e ictericia. Las pruebas bioquímicas mostraron unos niveles de ALT de 1117 UI/L y de bilirrubina de 7.6mg/dL junto con la presencia de anticuerpos anti-VHC y niveles séricos de ARN del VHC. Fue diagnosticado de hepatitis C aguda. Después de la normalización transitoria de la ALT y la desaparición del suero del ARN del VHC, los niveles de ALT se elevaron ligeramente de nuevo junto con la reaparición simultánea de ARN-VHC en noviembre de 2004 y se trató con interferón pegilado alfa2b durante 6 meses logrando una respuesta virológica sostenida. La investigación epidemiológica identificó que la transmisión del VHC se produjo de paciente a paciente durante el TAC (138). En este estudio se incluyó la muestra recogida en noviembre de 2004, antes de empezar el tratamiento con pegIFN+RBV.

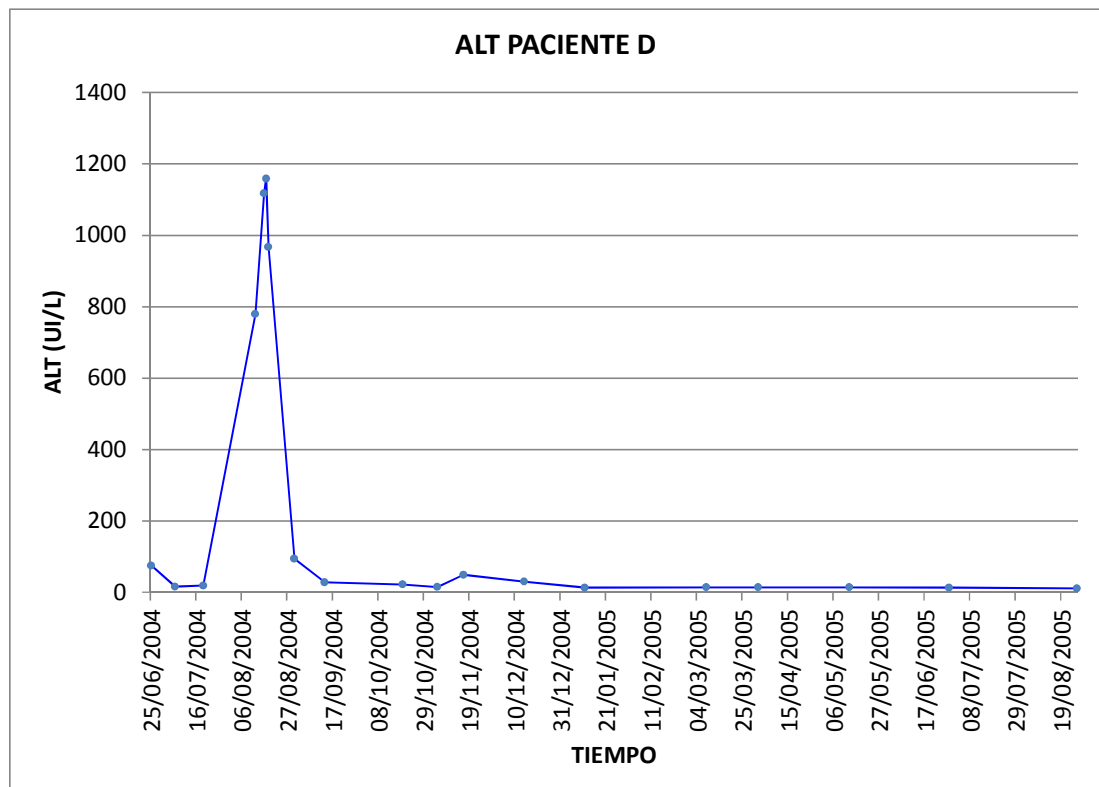
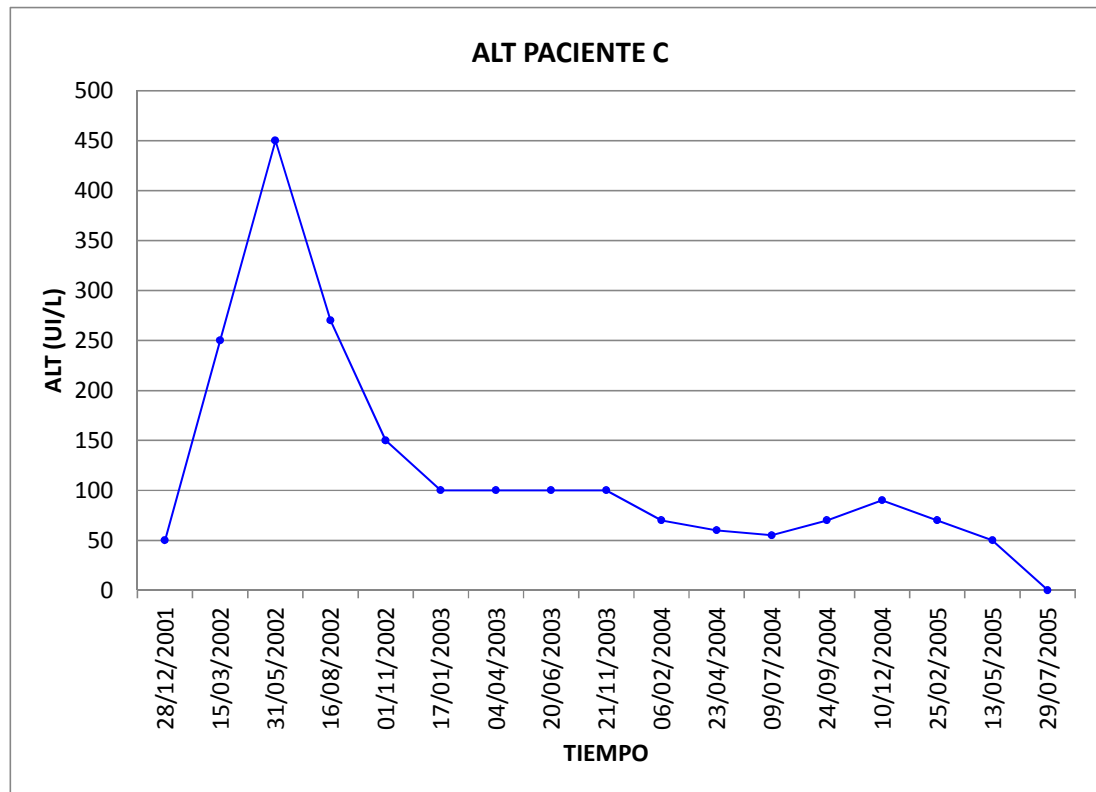


Figura 17. Valores de alaninaaminotransferasas (ALT) en pacientes C (gráfico superior) y D (gráfico inferior).

Paciente E trasplantado: El paciente E era un paciente varón de 69 años con infección crónica por VHC con genotipo 1b, que después de una cirrosis hepática, se sometió a un trasplante de hígado (la fecha del trasplante fue 21/06/2006). El nuevo hígado se re infectó, tal como ocurre en el 100% de los casos. Se estudió una muestra antes del trasplante (Epre) y otra a los 3 meses después del trasplante (Epost). Las pruebas bioquímicas mostraron unos niveles de ALT de 64 UI/L antes del trasplante hepático. Para facilitar la discusión, se consideraron las muestras Epre y Epost como si se tratasen de dos pacientes.

8.2 TÉCNICAS DE BIOLOGÍA MOLECULAR 1^{er} Y 2^o ESTUDIO

- *Conservación de la muestra*

La biopsia hepática se conservó en RNA later (Ambion Cambridgeshire, UK) durante 24 horas antes de su criopreservación a -80 ° C.

8.2.1 *Homogeneización de la muestra*

Se utilizó el Fast Prep (101 Savant), instrumento dotado de un rotor que se mueve a elevada velocidad y en múltiples direcciones, capaz de homogeneizar la muestra en un período de tiempo muy reducido, que depende de la dureza y cantidad de la muestra. La gran rapidez con la que el Fast Prep homogeneiza la muestra en presencia del buffer de extracción de ácidos nucleicos minimiza la degradación del ARN durante su manipulación.



Figura 19. Sistema de homogeneización de muestra Fast Prep

- Extracción de ARN de hígado

El ARN se extrajo con el kit Rneasy Mini Kit (Qiagen)

8.2.2 Control de cantidad y calidad de ARN extraído

Para el control de cantidad de ARN total extraído se utilizó un Nanodrop ND-1000, que es un espectofotómetro de UV/VIS de elevada precisión y sensibilidad, donde el volumen necesario para determinar la concentración de la muestra problema es de un 1µl. Este instrumento realiza un amplio barrido de espectro desde 220-750nm. La cantidad de ácidos nucleicos se mide a una longitud de onda de 260nm.

El control de calidad también se realizó mediante un equipo 2100 Bioanalyzer (Agilent Technologies), que permite medir la integridad del ARN utilizando de 20-250ng de muestra. Mediante una separación electroforética en unos microchips especializados, las diferentes muestras de ARN son separadas y monitorizadas (detección de fluorescencia con láser). El software del bioanalyzer genera un gel virtual que permite determinar el grado de integridad del ARN.

8.2.3 RT-PCR-NESTED

Se estudió toda la región NS3. El ARN fue convertido a cDNA mediante transcripción reversa con el uso de la Transcriptasa Reversa MMuLV con la mutación puntual H- (Promega). El “primer” empleado en la RT fue **A5319ds** ^{5'}**CAGGACGCCGCGCAACGAGCACCCA**^{3'}.

Tabla 11. Condiciones de la RT tanto de reactivos (tabla superior) como de temperaturas (tabla inferior)

Reactivos	Concentración en la reacción	Volumen (μl)
H ₂ O		10.3
Buffer Pfu 10x	1x	2
dNTP mix 10mM de cada uno/μl	2mM total	1
Primer 5319ds 20μM	20pmoles	1
Rnasin (Promega) 40U/μl	20U	0.5
Retrotranscriptasa MMuLV (Promega) 200u/μl	40U	0.2
ARN muestra problema		5
Volumen final		20

Pasos	Tiempo	Temperatura
	60 minutos	42°C
Retrotranscripción	5 minutos	99°C
	5 minutos	5°C
	α	4°C

Es interesante destacar, que se realizaron una serie de experimentos previos para poder realizar la copia del cDNA y la PCR en un mismo vial sin necesidad de precipitar entre las dos reacciones. Para ello, se usó el buffer de la Pfu para realizar la reacción de retrotranscripción y se pudo comprobar que la enzima MMuLV era capaz de tener muy buena actividad en este buffer (resultados no mostrados). Se preparó la mezcla de reacción de PCR siguiendo el esquema de la Tabla 3, usando el “primer” **A3349us** 5' **TGACATCATCAACGGCTTGCCTGTTTC** 3'. Se añadieron 80µl de la mezcla de PCR a cada vial correspondiente para conseguir un volumen final de reacción de PCR de 100µl.

Tabla 12. Condiciones de la PCR

Reactivos	Concentración reacción	Volumen (µl)
H ₂ O		70
Buffer Pfu 10x	1x	8
Primer A3349us 20 µM	20pmoles	1
Polimerasa Pfu turbo 2,5U/ µl	25U	1
cDNA		20
Volumen final		100

El cDNA se amplificó en el termociclador GeneAmp PCR 2720 mediante los siguientes ciclos de temperatura:

Tabla 13. Condiciones de Tª de la PCR

Pasos	Tiempo	Temperatura	Ciclos	
Desnaturalización	2 minutos	95°C		
	30 segundos	95°C		
	30 segundos	55°C	5x	
Extensión	2 minutos	72°C		
Desnaturalización	30 segundos	95°C		
Alineamiento	30 segundos	60°C	35x	
Extensión	2 minutos	72°C		Por
Extensión	10 minutos	72°C		

último, se realizó una Nested-PCR con una pareja de “primers” internos **A47us (up)** 5' **CGCCATCACGGCGTACGCCA3** 3' y **871ds (down)** 5' **GCATCCGTGGAGTGGCACTC** 3'.

Tabla 14. Condiciones de T°C de la Nested

Pasos	Tiempo	Temperatura	Ciclos
	2 minutos	95°C	
	30 segundos	95°C	
	30 segundos	60°C	30x
Extensión	2	72°C	
Extensión	10 minutos	72°C	
Conservación	α	4°C	

Para minimizar errores durante la amplificación, se usó la Pfu ADN polimerasa de alta fidelidad de copia (Stratagene, San Diego, CA,USA), que bajo nuestras condiciones experimentales, se ha demostrado que tiene una tasa de error entre 1.3×10^{-6} (139) y

6.5×10^{-7} (140). Después de 70 ciclos de amplificación, el número esperado de mutaciones según el fabricante es 1 error en 9.1×10^5 nucleótidos secuenciados. En nuestro estudio control uno de los clones (1893 nucleótidos) de secuencia conocida, fue sujeto a 70 ciclos adicionales de PCR en 20 amplificaciones independientes. No se detectaron mutaciones después del análisis de 8140 nucleótidos. Estos resultados demuestran que la heterogeneidad observada es independiente de cualquier artefacto durante el procedimiento de amplificación (141).

Clonación

- Determinación del tamaño molecular del producto amplificado

Los segmentos amplificados producto de la Nested se visualizaron por electroforesis en gel de agarosa al 1.5% y se tiñeron con bromuro de etidio.

- Purificación del ADN

La purificación del ADN amplificado resultante de la Nested se realizó lo que se utilizó el kit de purificación en gel de agarosa de QIAquick Gel Extraction Kit (Qiagen, GMBH, Hilden).

- Reacción de clonación

El producto purificado de la PCR-Nested con extremos romos se clonó en un vector plasmídico (pCR 4Blunt-TOPO) y éste a su vez, se transformó en células *E.coli* competentes TOP10. Esta clonación se realizó con el Kit Zero Blunt TOPO PCR Cloning Kit for Sequencing (Invitrogen), el cual permite la inserción directa en el plásmido de productos de PCR “blunt-end” (extremos romos) que se generan por la polimerasa Pfu turbo High fidelity para su posterior secuenciación.

- pCR 4Blunt-TOPO

El plásmido pCR 4Blunt-TOPO tiene un genoma compuesto por 3956 nucleótidos, y posee un gen que le confiere resistencia a la ampicilina, permitiendo

sólo el crecimiento de los recombinantes positivos en placas de Petri con medio LB con ampicilina, después de la transformación en células *E.coli*.

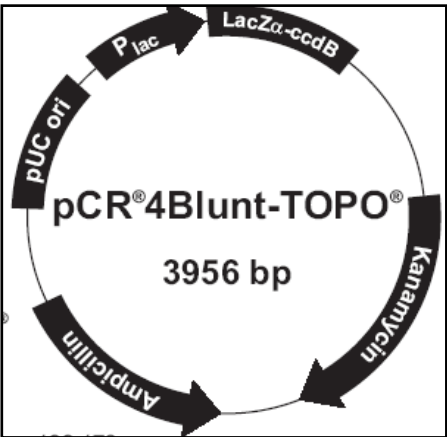


Figura 20. Mapa plásmido pCR-Blunt-TOPO

Los reactivos y las cantidades empleadas en la mezcla de clonación fueron los siguientes:

Tabla 15. Reactivos de la mezcla de clonación.

Reactivos	Control + (μl)	Control – (μl)	Producto de PCRNESTED (μl)
Volumen DNA	0	0	1
Sales	1	1	1
H ₂ O	4	4	1
Vector plasmídico	pUC 19 Control	pCR 4Blunt-TOPO	pCR 4Blunt-TOPO
(10ng/μl)	1	1	1
Volumen final	6	6	6

- Se mezcló la reacción suavemente.
- Se incubó 30 minutos a RT. En el protocolo del Kit indica que la incubación con la mezcla de clonación puede ser en un intervalo de duración de 30 segundos hasta 30 minutos, decantándonos por este último, con la finalidad del aumento del rendimiento final de transformantes que contengan el inserto.
 - Se añadieron 2µl (como indica el protocolo del kit) de la reacción de clonación del control negativo y positivo a los viales que contenían las bacterias químicamente competentes TOP10 y se mezcló suavemente. Para la clonación de la Nested-PCR de NS3 del VHC se añadieron 2µl y 4µl de la mezcla de reacción de clonación, con esta última cantidad se obtuvo un mayor número de colonias después de la siembra de las bacterias *E. coli* competentes transformadas con el inserto de serin-proteasa NS3.
 - La reacción se incubó en hielo durante 30 minutos que es el tiempo máximo recomendado y que hemos comprobado que incrementa la eficiencia de transformación.
- [Transformación en *E. coli* TOP10 competentes](#)
 - Las bacterias químicamente competentes se transformaron con el plásmido mediante un choque térmico de las células durante 30 segundos a 42°C en el baño termostático, sin agitación y se transfirieron los tubos al hielo.
 - A cada vial se añadieron 250µl de medio S.O.C. a RT, medio rico en sales, cuya función es estabilizar las membranas de las bacterias químicamente competentes.
- [Análisis y selección de colonias](#)
 - Se agitaron los tubos horizontalmente en un agitador a 200 rpm a 37 ° C durante 1 hora.
 - Las placas de Petri con el medio de LB+Ampicilina (100µg/ml) se precalentaron en la estufa durante media hora antes de sembrar las muestras.

- Se sembraron 50µl, 100µl y 150µl de cada muestra en las placas de LB y se incubaron a 37 ° C durante 24 horas. En las placas donde se obtuvo un mayor número de colonias fueron aquellas donde se sembró un volumen de 150µl, pero la incubación se realizó entre 18-20 horas, para evitar el solapamiento de las colonias y aparición de satélites.

- [Purificación del ADN clonado en *E. coli* competentes](#)

El conjunto de colonias obtenidas resistentes a Ampicilina, se picaron con agujas estériles y se introdujeron en placas microtiter de 96 pocillos, donde en cada uno se había depositado previamente 200µl de medio LB líquido con ampicilina para su conservación. Cada colonia picada se pasó a un pocillo diferente de la placa de microtiter. Dicha placa se incubó 24 horas a 37°C. Las colonias obtenidas se pasaron de los pocillos de la placa de microtiter a tubos de vidrio, que contenían 1ml de medio LB con ampicilina, mediante puntas de pipeta estériles. Una vez efectuada esta operación, se dejaron incubando 24 horas a 37°C en el agitador a 230rpm. Una vez crecieron las bacterias, se pasaron de los tubos de vidrio a eppendorf de 1.5ml y se centrifugaron durante 5 minutos a 13.000rpm y RT.

Las colonias obtenidas de las clonaciones, se sometieron a una lisis con el Kit *QIAprep Spin Miniprep Kit Using a Microcentrifuge (Qiagen)* para la obtención del ADN, y su posterior secuenciación. Este método está basado en una lisis alcalina de las bacterias, seguida de una adsorción del ADN en una membrana de sílice en presencia de una solución salina.

- [Secuenciación del ADN obtenido de los clones de la región NS3 VHC](#)

Los productos de amplificación se secuenciaron utilizando terminadores de reacción marcados con fluorescencia (“dye terminators”) de la casa Applied Biosystems. Para ello se empleó el Kit de secuenciación ABI Prism BIGDYE Terminator versión 1.1. La secuenciación automática fue llevada a cabo inicialmente en nuestro laboratorio con el secuenciador ABI PRISM 310 y posteriormente se utilizó el servicio de secuenciación ABI PRISM 310 de 16

capilares de la Unidad Científico-Técnica de Soporte (UCTS) del Institut de Recerca del Hospital Universitario Valle de Hebrón.

Las secuencias obtenidas se procesaron con el programa SeqScape v. 2.0 de Applied Biosystems. El análisis y alineación a nivel de nucleótidos y aminoácidos se realizó con el programa GeneDoc.

8.3 TÉCNICAS DE BIOLOGÍA MOLECULAR 3^{er} ESTUDIO

8.3.1 IL28 y HLA

Para los estudios genómicos se obtuvo la aprobación del CEIC y de los pacientes. El ADN genómico fue extraído con el kit comercial de ADN Flexigen (Qiagen, Hilden, Alemania). El kit IL28B LightMix (TIB MolBio) se usó para definir el genotipo del polimorfismo bialélico, rs12979860 (CC, CT, TT), fuertemente asociado con respuesta al tratamiento. El kit contiene cebadores para amplificar una región de 139pb y sondas específicas para el alelo -3176C. A partir de 5µl de ADN genómico, cada alelo del polimorfismo fue identificado por curvas de melting (alelo T $T_m=51^{\circ}\text{C}$ y alelo C con una $T_m=59^{\circ}\text{C}$), usando el LightCycler 2.0 (Roche Diagnostics GMBH, Mannheim, Alemania). El haplotipo de HLA se determinó mediante secuenciación por Sanger.

8.3.2 Antígenos del VHC

Los pools de péptidos derivados de secuencias de genotipo 1a que cubrían la región completa de NS3 se adquirieron de Beiresources (Biodefense and emerging infectious research resources repository beiresources Manassas VA USA) y 63 péptidos de 20 aminoácidos que cubrían también la región NS3 sintetizados por CIMA (Universidad de Navarra, Pamplona.)

8.3.3 *Ensayo de Enzyme-Linked Immunospot (ELISpot)*

Para el ensayo de IFN- γ ELISpot se realizaron duplicados de 1×10^5 células CD4 basales y de células CD4 expandidas que fueron cultivadas a 37° C en una proporción de 1:4 con células presentadoras de antígeno (CPA) autólogas irradiadas (30 Gy) pulsadas con diferentes pools de péptidos o controles. Los pools para la proteasa NS3 fueron: 1027-1117 y 1117-1217 y para la NS3- helicasa fueron 1227-1317, 1327-1417, 1427-1517 y 1527-1642. El ELISpot se desarrolló siguiendo el protocolo siguiente: las células se cultivaron en placas de 96 pocillos de fondo redondo pre-incubadas con anticuerpo monoclonal anti-IFN- γ (1-D1K, Mabtech, Suecia) y con una concentración de 2 μ g/ml con los diferentes antígenos del VHC, 2 μ g/ml de PHA (Fitohemaglutinina) o de buffer control. Después de 48 horas a 37°C y 5% de CO₂, las células se eliminaron mediante lavado de las placas 5 veces con PBS. Se añadieron 100 μ l de fosfatasa alcalina conjugada con anticuerpo monoclonal (7-B6-1, Mabtech, Suecia) a cada pocillo a una concentración final de 1 μ g/ml y se incubó a 2h a temperatura ambiente. El conjugado no unido se eliminó con lavados de las placas con PBS y se desarrolló una incubación durante 5 minutos con una solución de sustrato de 5-bromo-4-cloro-3-indolil fosfato/nitroblue tetrazolium (BCIP/NBT-plus; Mabtech, Suecia). La reacción de color se detuvo mediante lavado con agua del grifo. Las células formadoras de puntos péptido-específicas (spot-forming cells-SFC-) se contaron con un lector de ELISpot (Auntimmun Diagnostika GMBH, Strassberg, Germany) con la ayuda del software (ELISpot 4.1 SR program). Los resultados se expresaron como el número de SFC productoras de IFN- γ por 10^5 células CD4 (SFC IFN- γ /10⁵ CD4). Los ensayos con ruido de fondo alto (media de 40 SFC/pocillo con buffer control) o con nula respuesta a PHA se excluyeron. Para normalizar el número de células específicas de VHC productoras de IFN- γ , se descontó el número de SFC productoras de IFN- γ sin péptidos del número de SFC con presencia de péptidos.

8.3.4 Análisis de la región NS3 mediante secuenciación masiva

Se realizó una secuenciación masiva (Ultra-deep pyrosequencing-UDPS-) mediante reactivos de la serie GS-FLX Titanium en la plataforma Roche 454 GS-FLX) de 7 fragmentos (amplicones) de 400 nt correspondientes a la región no estructural NS3 (proteasa y helicasa) en los pacientes A, B (infectados por el VHC genotipo 1a); C y D (infectados por el VHC genotipo 1b); E-pre y E-post (genotipo 1b). Para la amplificación de la proteína NS3 se utilizaron cebadores específicos, y teniendo en cuenta que la química de GS-FLX Titanium permite la secuenciación de 400 nucleótidos, se dividió la NS3 en 7 amplicones solapados. Los amplicones fueron enumerados del 1 al 7 desde el extremo 5' hasta el final de la región codificante 3' de NS3. Cada amplicón se analizó independientemente. Para minimizar los errores de muestreo y amplificación, se realizaron RT-PCRs-Nested por triplicado. Finalmente, los productos de PCR se cuantificaron, se mezclaron y purificaron en gel de agarosa mediante electroforesis, previo a la secuenciación masiva por ultra-deep-pyrosequencing (UDPS) plataforma 454/Roche (Roche, Brandford, CT 06405, USA). Se realizaron 3 placas 454 GS FLX Titanium, y cada una se dividió en dos carriles (en cada carril se depositó la muestra de un paciente). El sistema generó un archivo en FASTA con los “reads” secuenciados de cada carril.



Figura 21. Aparato 454 GS FLX Titanium

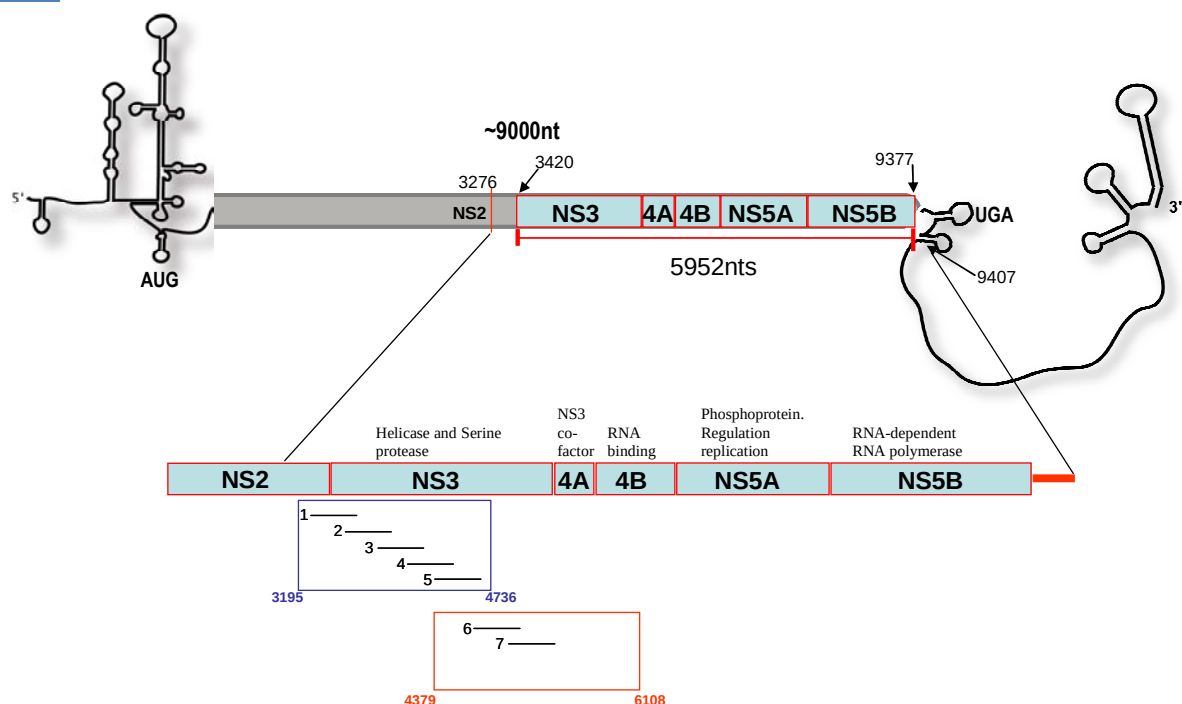


Figura 22. Esquema de la región no estructural del VHC donde se representan los 7 amplicones de 400 nts cada uno que cubrían la proteína NS3.

Nombre de Cebadores		Secuencias	Tamaño amplicones
RT-PCR1	NS2up315pcr	CGTATCGCCTCCCTCGCGCCATCAGCAYCTBACYCCNCTBCRRGAYTGGGC	1542pb
	NS3d4736	CGTATCGCCTCCCTCGCGCCATCAGCAYCTBACYCCNCTBCRRGAYTGGGC	
Nested1	NS2up3272	CGTATCGCCTCCCTCGCGCCATCAGCMVVATGAGACCAAGCTCATYAC	442pb
	NS3d3663	CTATGCGCCTTGCCAGCCCGCTCAGGRTCTGTCTYACATTRGTRTACATYTG	
Nested2	NS3up3531	CGTATCGCCTCCCTCGCGCCATCAGACYGCDRCVCARWCYTTCTTRGC	489pb
	NS3d3969	CTATGCGCCTTGCCAGCCCGCTCAGAVACCGGRGAYCKCATRGTGTYTC	
Nested3	NS3up3816	CGTATCGCCTCCCTCGCGCCATCAGTCTYACTYGAARGGYTCBTCVGG	457pb
	NS3d4222	CTATGCGCCTTGCCAGCCCGCTCAGGARTAYGTRATRGGRWYRCCNGTGGT	
Nested4	NS3up4038	CGTATCGCCTCCCTCGCGCCATCAGGGHAGCGGYAARAGBACYAARGT	424pb
	NS3d 4411	CTATGCGCCTTGCCAGCCCGCTCAGACRGTRRYNGAKCCCGGAGG	
Nested5	NS3up4279	CGTATCGCCTCCCTCGCGCCATCAGTAATNTGYGAYGARTGCCAYTC	456pb
	NS3d4684	CTATGCGCCTTGCCAGCCCGCTCAGCMAGRTGGTCTYGVCANACGG	
RT-PCR2	NS3up4379	CGTATCGCCTCCCTCGCGCCATCAGCGCCACYGCIACNCTCCGGG	1730pb
	NS4Bd6108	CTATGCGCCTTGCCAGCCCGCTCAGCGAAWGCTATYAGCCGGTTCATCC	
Nested6	NS3up4588	CGTATCGCCTCCCTCGCGCCATCAGCNTAYTACCGVGBCTYGA YGTRTC	470pb
	NS3d5007	CTATGCGCCTTGCCAGCCCGCTCAGCMAGRTGGTCTYGVCANACGG	
Nested7	NS3up4905	CGTATCGCCTCCCTCGCGCCATCAGGAYGCRGGCTGYGCTTGTA	447pb
	NS3d5301	CTATGCGCCTTGCCAGCCCGCTCAGCNAGGTCRGYGMCATGCAT	

Tabla 16. Secuencias de cebadores para la amplificación de NS3.

- Control de calidad de los amplicones

El control de calidad se realizó mediante el equipo 2100 Bioanalyzer (Agilent Technologies), que permite obtener una inspección visual de la integridad del ARN utilizando de 20-250ng de muestra. Mediante una separación electroforética en unos microchips especializados, las diferentes muestras de ARN son separadas y monitorizadas (detección de fluorescencia con láser). El software del bioanalyzer genera un gel virtual que permite determinar el grado de integridad del ARN.

La concentración de ADN se determinó mediante el uso de un colorante fluorescente que se une al ADN de doble cadena (PicoGreen, Molecular Probes), mediante el Kit PicoGreen dsDNA Quatitation Kit (Invitrogen, Carlsbad, California), que tiene un máximo de excitación a 480nm y un pico máximo de emisión a530nm.

- Diseño de los cebadores

La preparación de las muestras de ADN para la secuenciación de amplicones consiste en una simple reacción de amplificación por PCR con cebadores especiales (cebadores de fusión). Los cebadores de fusión permiten la unión de los fragmentos de ADN a las bolas de captura de ADN y su posterior amplificación clonal mediante una PCR en emulsión (emPCR). Poseen un tamaño entre 20-25pb. Las últimas 4 bases de los cebadores de fusión es la secuencia clase "TCAG".

Forward primer (Primer A-Key):

5'-CGTATCGCCTCCCTCGCGCCA**TCAG**-template specific sequence-3'

Reverse primer (Primer B-Key):

3'-CTATGCGCCTTGCCAGCCCGCT**TCAG**-template specific sequence-5'

Figura 23. Estructura de los cebadores de fusión A y B.

Para la realización de la emPCR fueron necesarios dos kits (Figura 24): uno para la secuenciación de la librería del primer A (GS emPCR Kit II) y el segundo para la secuenciación del Primer B (GS emPCR Kit III). Es decir, para la emPCR de cada amplicón se utilizaron 2 cebadores (A y B).

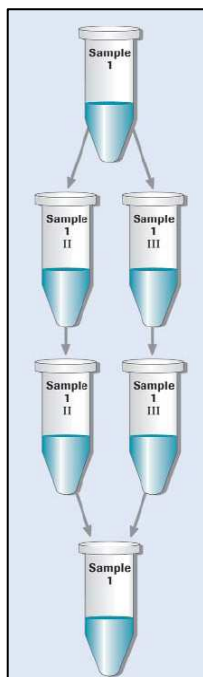


Figura 24. GS emPCR kit II y kit III.

- [emPCR](#)

La librería sstDNA de amplicones se consiguió diseñando los cebadores con los adaptadores A y B en cada extremo, tal y como se especifica en la página 135 Figura 24. Se mezclan cada bola con las moléculas de ADN, de manera que cada bola, en condiciones óptimas, contendrá una única molécula de sstDNA de la librería. La bola unida a la cadena de ADN se emulsiona con los reactivos de amplificación en una micela de agua y aceite. Cada perla queda englobada en su propio microreactor dentro del cual ocurre la amplificación mediante PCR. El resultado será una bead inmovilizada conteniendo fragmentos de ADN amplificados clonalmente.

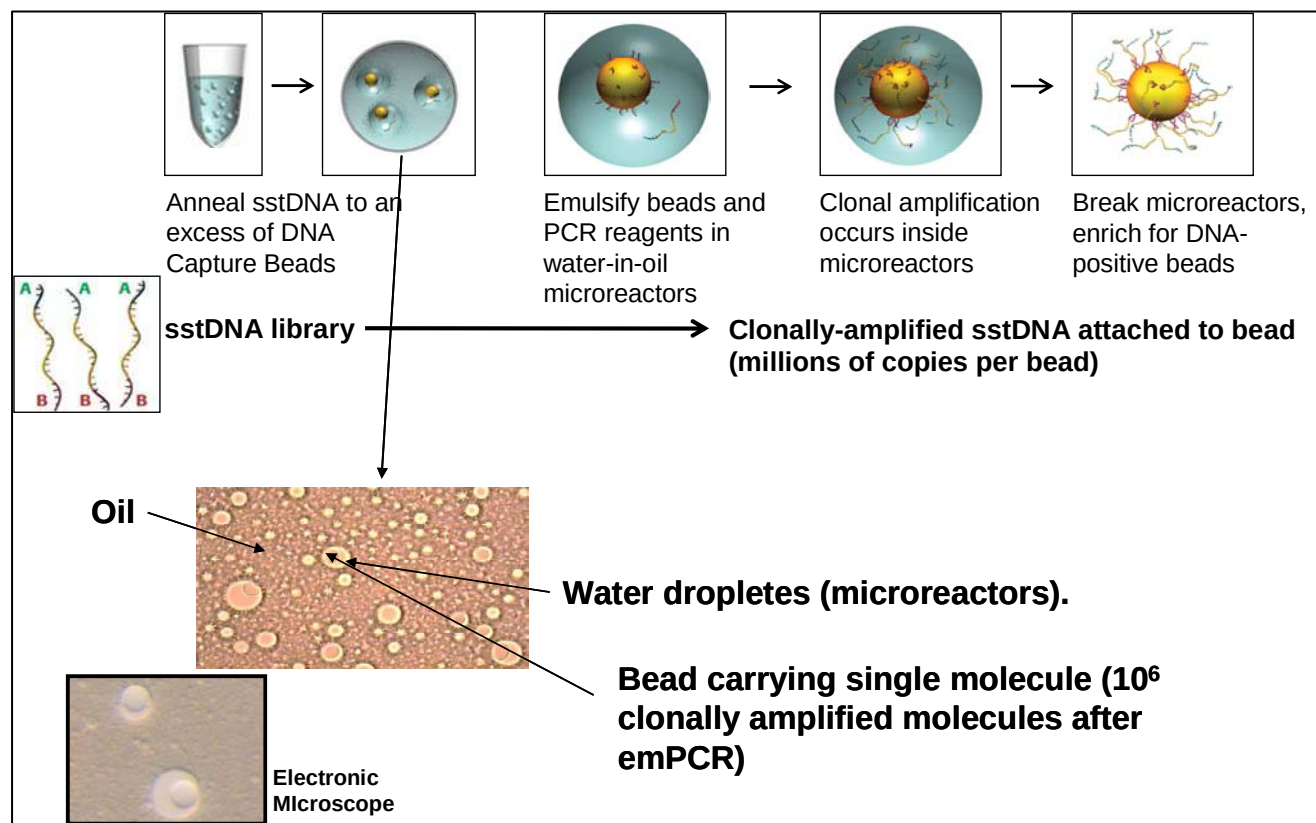


Figura 25. Pasos secuenciales de la emPCR.

- Secuenciación

Las beads unidas a los fragmentos de la librería ADN_{sst} se añaden al DNA Bead Incubation Mix (contiene ADN polimerasa) y se distribuyen en capas dentro de la placa junto con unas beads enzimáticas que contienen luciferasa y sulfurilasa. La capa de beads enzimáticas asegura que las perlas con ADN permanezcan en el interior del pocillo durante la reacción de secuenciación. El proceso de deposición en la placa maximiza el número de pocillos que contienen una única perla con ADN amplificado (evitando más de una bead unida a ADN_{sst} por pocillo). Cuando la placa está correctamente cargada se coloca en el equipo donde los reactivos de secuenciación (tampones y nucleótidos) fluirán a través de la placa. Durante el flujo de nucleótidos, cada una de los cientos de miles de beads con millones de copias de ADN se secuencian en paralelo. Si un nucleótido es complementario a la cadena molde en algún pocillo, la polimerasa extiende la hebra existente de ADN mediante la adición de nucleótido(s). La adición de cada nucleótido genera por reacción la

emisión de un fotón. La señal lumínica es recogida por la cámara CCD del equipo. La intensidad de la señal es proporcional al número de nucleótidos incorporados en un solo flujo de nucleótidos en el respectivo pocillo.

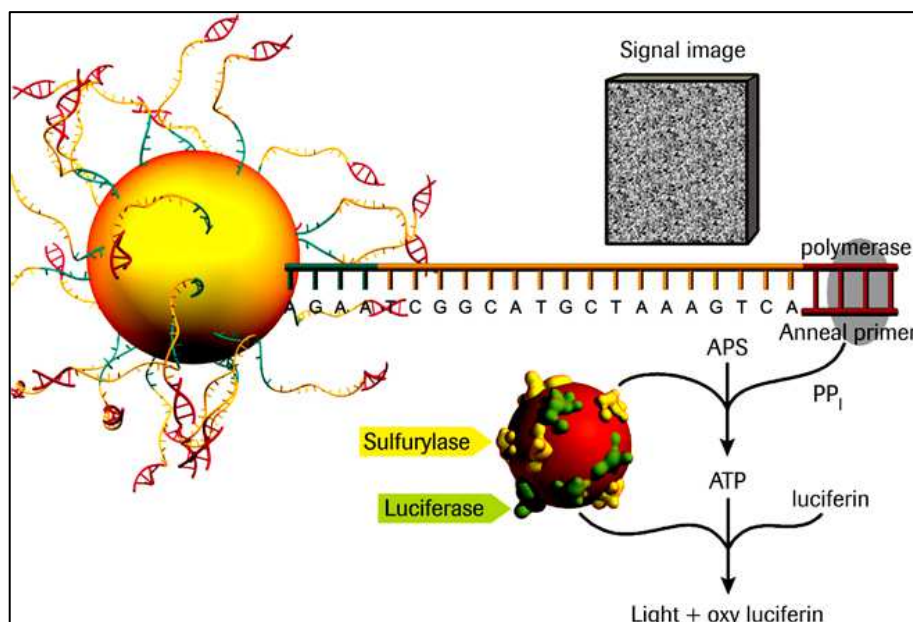


Figura 26. Reacción de secuenciación.

8.3.5 ALGORITMO PARA INTERPRETAR LAS SECUENCIAS

Las lecturas (“reads”) correspondientes a cada amplicón se identificaron por su correspondiente primer. Los datos se validaron de acuerdo a procedimientos ya publicados (Homs M. et al. Nucleic Acids Res. 39:8457-8471; Campbell P.J. et al. Proc. Natl. Acad. Sci. U.S. A 105:13081-13086; Huse S.M. et al. Genome Biol. 8:R143; Zagordi O. et al. Nucleic Acids Res. 38:7400-7409). Como filtros de calidad, se excluyeron las secuencias que:

- Tenían más de 2 errores con respecto al cebador.
- Las que no cubrían todo el amplicón
- Las que tenían Ns en la secuencia
- Las que tenían inserciones o deleciones con respecto a la secuencia de referencia.

Las sustituciones con respecto a la secuencia de referencia fueron evaluadas estadísticamente por un filtro de Poisson, utilizando una matriz de sustituciones de error empírica obtenida por secuencias clonales adicionales obtenidas por UDPS. Una vez identificadas y filtradas los “reads” se colapsaron en haplotipos y se determinaron sus correspondientes frecuencias. Los haplotipos con un solo “read” se rechazaron. Los haplotipos resultantes con sus respectivas frecuencias son considerados como una muestra representativa de la composición de la población viral, y se utilizaron en los análisis siguientes, incluyendo un análisis de las mutaciones posición por posición y la evaluación de la complejidad mediante “sliding windows”, a nivel de nucleótidos y a nivel de aminoácidos.

8.3.6 ANÁLISIS DE LAS SECUENCIAS

Las definiciones de los términos utilizados en el análisis de secuencias son las siguientes:

Amplicón: un fragmento de secuencia de ARN o ADN que es amplificada mediante cebadores específicos.

Read: cada una de las secuencias identificadas a través de síntesis durante el proceso de la UDPS, correspondiente a un amplicón.

Mutación: sustitución de un nucleótido por otro nucleótido. Cuando se habla del número total de mutaciones se refiere al número total de mutaciones diferentes, independientemente de sus frecuencias dentro de la población. La mutación se entiende como un cambio respecto al haplotipo dominante.

Haplotipo: en el contexto del análisis de los amplicones, un haplotipo es cada una de las diferentes secuencias identificadas por UDPS, cada una de ellas representada por un número de reads que debe corresponder aproximadamente a su frecuencia en la población de la quasiespecie obtenida por muestreo. Los haplotipos son entonces el conjunto de secuencias únicas en el amplicón dentro del conjunto de la población viral.

El análisis se realiza posición por posición, en el que se obtiene una lista de todos los sitios polimórficos, junto con las frecuencias de mutación de nucleótidos correspondientes. Se analiza cada amplicón según una heurística que favorece las secuencias con un mayor número de lecturas. El recorte de los reads para conectar los amplicones es un paso importante en el flujo de trabajo.

Análisis de nucleótidos de cada amplicón

1. **Filtrado de reads**: se realizó un filtrado de reads de manera que cada haplotipo estuviera representado por un mínimo de lecturas, excluyendo todos los haplotipos con menos de 2 reads. Esto contribuyó a aumentar el peso de la evidencia y evitar posibles artefactos.
2. **Complejidad de quasiespecies**: la complejidad de las quasiespecies se caracterizó por una serie de parámetros para realizar estudios comparativos entre poblaciones virales, que se describen a continuación:

Diversidad nucleotídica (p_i, π): tiene en cuenta el número medio de mutaciones entre cada par de secuencias en la población viral.

Frecuencia de mutación (M_f): obtenida como la relación del número de mutaciones observadas por el número total de nucleótidos secuenciados. El número de nucleótidos secuenciados es el producto del número de reads por el tamaño de cada read. El cálculo de las mutaciones observadas tiene en cuenta las frecuencias de los haplotipos dentro de la población viral, y el número de mutaciones en cada haplotipo.

Entropía de Shannon (H): la entropía de Shannon mide la variabilidad de las quasiespecies teniendo en cuenta el número de haplotipos y sus frecuencias:

$$H = - \sum_i p_i \ln(p_i)$$

Donde p_i es la frecuencia relativa del haplotipo i -ésimo dentro de la población viral, y la suma se extiende a todos los haplotipos. Un valor alto de H implica una variabilidad alta. El mínimo valor de entropía, es 0, cuando tenemos un haplotipo único.

Entropía de Shannon normalizada: es una medida de diversidad de haplotipos, independientemente de sus frecuencias de mutación.

$$S_n = - \sum_i p_i \ln(p_i) / \ln(h)$$

Al dividir H por el logaritmo del número de haplotipos S_n varía entre 0 y 1, con un valor de S_n , de 1, como la entropía máxima.

Gráficos de barras: se representaron gráficamente los sitios polimórficos (eje de ordenadas) con las frecuencias de mutación relativas respectivas (eje abscisas). En la parte inferior se da una representación reducida en color que muestra la secuencia consenso, y las barras apiladas sobre ella dan el porcentaje de mutación para cada nucleótido dentro de la población. Cada nucleótido está representado por un color diferente. La línea en 0 diferencia, por debajo, el nucleótido mayoritario, y por encima, la mutación.

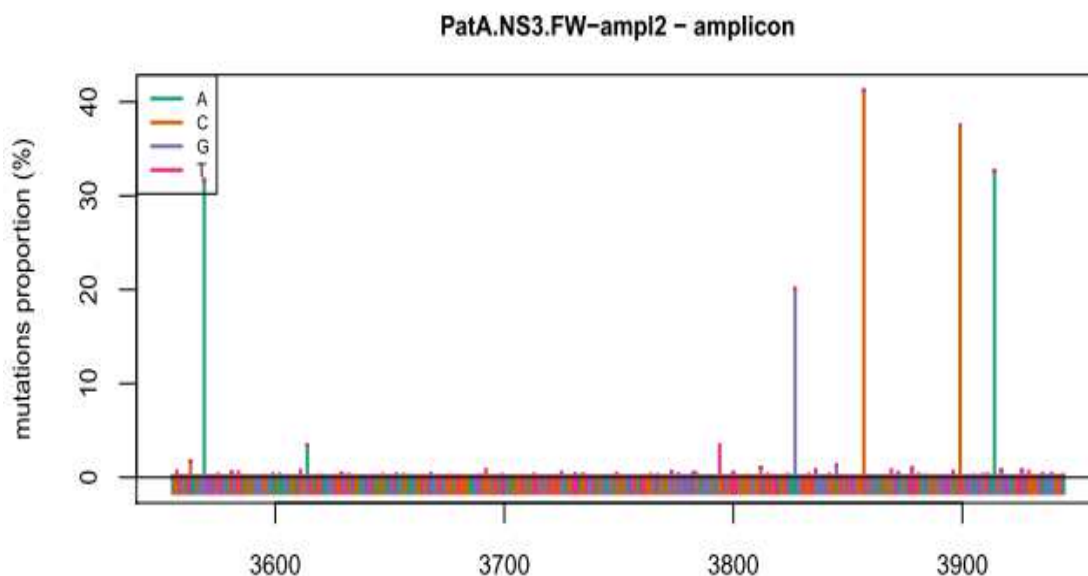


Figura 27. Ejemplo de gráfico de barras con las frecuencias de mutación relativas.

- Contenido de información de una posición (ic_i)

Dada la frecuencia relativa en la población de cada nucleótido p_i , en un sitio determinado, el contenido de información, en unidades de bits se obtiene como:

$$ic = \log_2(4) + \sum_i p_i \log_2(p_i) = 2 + \sum_i p_i \log_2(p_i)$$

La entropía es una medida de diversidad. Aquí esta medida se invierte para proporcionar el contenido de información de un sitio dentro de una población. Ambos conceptos están inversamente relacionados. Si la entropía de Shannon de las quasiespecies es una entropía que mide por filas (secuencias), el contenido de información en un sitio es un análisis de la entropía por columnas (sitios).

Un sitio conservado 100% tiene 2 bits de información, mientras que un sitio con 25% de cada nucleótido mostrará un contenido de información de bits 0.

Gráficos de densidad mutacional (Montserrat plots)

Los Montserrat plots son gráficos que sirven para visualizar la diversidad de la población respecto al haplotipo dominante. Tomando como referencia la secuencia dominante de la población viral en cada amplicón, este gráfico representa la densidad del número de mutaciones en la población viral respecto a esta secuencia de referencia. El histograma habitual se transforma en un gráfico de densidad usando a kernel gaussiano con un ancho de banda de 0,3, transformando el gráfico en continuo y virtualmente más cercano a la dinámica de la quasiespecie. Se superpone la densidad que se deriva, en las mismas condiciones, de una distribución de Poisson, cuyo parámetro es el número medio de mutaciones. Cuando ambas curvas coinciden la población tiende a ser más aleatoria, mientras que cuando divergen notablemente, la población está más sujeta a factores evolutivos, representando eventualmente 2 o más subpoblaciones separadas por un cierto número de mutaciones.

Análisis de la región de nucleótidos

En general, la región de interés en el genoma viral se compone de una serie de diferentes amplicones, por lo general cubren unos cuantos miles de bases. El solapamiento entre amplicones se resuelve apoyándose en el amplicón con mayor número de reads.

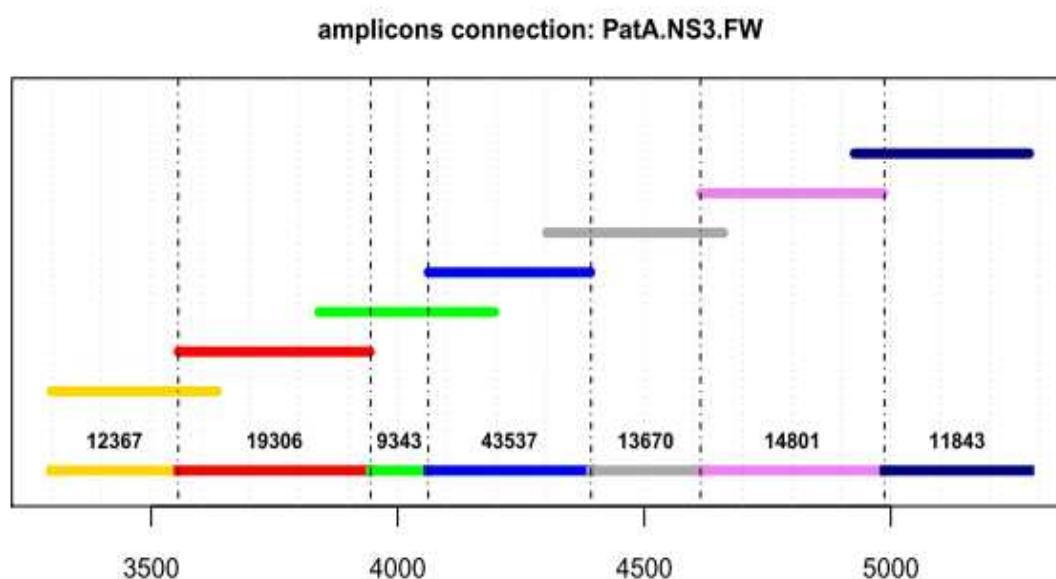


Figura 28. Ejemplo de solapamientos de los 7 amplicones de NS3 del paciente A

El análisis de la región se ha realizado de forma independiente para las cadenas forward y reverse, y el siguiente paso fue combinarlas para obtener una imagen final posición por posición de la región de interés. Se consideraron primero los sitios polimórficos comunes, testando la homogeneidad de la frecuencia de sustituciones de nucleótidos entre ambas cadenas mediante el test de Fisher. Los p-valores resultantes se ajustan por la corrección Benjamini y Hochberg. Los sitios no significativamente diferentes son promediados para las cadenas forward y reverse, mediante el correspondiente número de reads de cada cadena. Para los sitios significativamente diferentes, si el número de reads es muy desequilibrado entre ambas cadenas, se toma la frecuencia de mutación dada por el número de secuencias más abundantes. Al final tenemos un cuadro con las frecuencias de cada nucleótido en cada sitio polimórfico que finalmente es aceptado. El informe representa los resultados obtenidos en cada caso, dando para cada sitio polimórfico aceptado, las coordenadas, las frecuencias de nucleótidos en %, y el p-valor ajustado por un test múltiple.

Chi-square test

Common positions, non-significantly different,
with BH adjusted p-values above 0.05
Giving the average mutation frequencies weighted by
reads number.

	pA	pC	pG	pT	adj.p.val
3317	99.218	0.000	0.397	0.385	0.461
3320	0.000	99.261	0.000	0.739	0.319
3323	0.000	0.133	0.000	99.867	0.258
3326	0.000	99.620	0.000	0.380	0.138
3329	0.000	99.833	0.000	0.167	0.442
3337	0.128	0.000	99.872	0.000	0.755
3338	0.000	99.561	0.000	0.439	0.557
3339	0.000	66.115	0.000	33.885	0.189
3350	0.000	0.249	0.000	99.751	0.325
3356	0.000	99.779	0.000	0.221	0.557
=====					
5255	0.142	0.000	99.858	0.000	0.964
5264	0.000	99.854	0.000	0.146	0.609
5267	0.000	99.817	0.000	0.183	0.930
5270	0.301	0.000	99.699	0.000	0.995

Common positions, significantly different
with BH adjusted p-values below 0.05

If the number of reads is markedly unbalanced, the mutation in the most abundant strand, the minimum frequency otherwise.

	pA	pC	pG	pT	adj.p.val
3305	0.384	0.000	99.616	0.000	0.020
3314	0.000	99.241	0.000	0.759	0.000
3347	0.000	0.982	0.000	99.018	0.001
3524	0.000	98.932	0.000	1.068	0.027
3545	0.000	67.404	0.000	32.596	0.000
3563	0.000	1.343	0.000	98.657	0.020
=====					
4868	0.000	99.662	0.000	0.338	0.009
4886	0.000	99.480	0.000	0.520	0.005
5138	0.000	99.820	0.000	0.180	0.020

Summary of total mutation frequency in the forward strand only:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.1010	0.1070	0.1290	0.1405	0.1420	0.4251

No significative mutations in the FW strand only.

Summary of total mutation frequency in the reverse strand only:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.1000	0.1090	0.1330	0.1417	0.1655	0.3010

RV only, significative mutations to the test,
 RV/FW reads ratio higher than 1.5
 Total mutation frequency higher than 0.1 %
 Reads number higher than 5000
 Adjusted p-values below 0.05

	pA	pC	pG	pT	p.vals	adj.p.val
reads						
3297	0.000	0.000	0.205	99.790	1.904e-11	4.251e-10
22940						
3311	0.000	99.820	0.000	0.179	4.111e-10	3.935e-09
22940						
3315	0.100	0.000	99.900	0.000	4.465e-06	7.671e-06
22940						
3333	99.900	0.000	0.100	0.000	4.465e-06	7.671e-06
22940						
=====						
3533	0.000	0.135	0.000	99.860	7.047e-08	2.623e-07
22940						
3983	0.000	99.890	0.000	0.105	6.124e-05	8.373e-05
17210						
4019	0.000	99.770	0.000	0.232	6.828e-10	5.147e-09
17210						

Para las mutaciones puntuales se requiere un umbral de frecuencia mínima, el valor predeterminado se establece en el 0,1%. Para los sitios polimórficos en una sola cadena se requiere, además, una frecuencia mínima de mutación total y un número mínimo de reads. En la actualidad los valores por defecto son de un p-valor ajustado de 0,05, un porcentaje de 0,1%, como frecuencia mínima de mutación total, y un mínimo de 5000 reads.

RATIO dN/dS

Se calculó la relación de dN/dS como la relación entre los valores medios de dN (mutaciones no sinónimas por posiciones no sinónimas) y dS (mutaciones sinónimas por posiciones sinónimas) que resultan al comparar pares de reads en la población viral. Las presiones evolutivas de las proteínas son cuantificadas por la tasa de sustitución en posiciones sinónimas y no sinónimas. Como consecuencia de la degeneración del código genético, el resultado esperado de mutaciones al azar no sinónimas se aproxima a 75% y las mutaciones sinónimas a un 25%. La relación dN/dS se desarrolló originalmente para el análisis de secuencias genéticas de especies divergentes. Sin embargo, la medida de dN/dS también se suele aplicar a las muestras de una sola población. En el caso de las poblaciones de virus, entre las secuencias no existe fijación a lo largo de linajes independientes, sino más bien la segregación de polimorfismos en una sola población. Por lo tanto, la relación entre dN/dS es cuantitativamente diferente en las muestras de una sola población en comparación con las muestras de linajes divergentes. En este último caso, la relación dN/dS=1 se interpreta como neutralidad, cuando dN/dS<1 como selección negativa y dN/dS>1 como selección positiva. Pero en el caso de las poblaciones de virus (estudio de muestras dentro de una misma población) cuando dN/dS es mayor o menor que 1, se interpreta que existe presión evolutiva, pero no se puede afirmar si hay selección positiva o negativa (142). En nuestro estudio se calculó, a partir de los haplotipos observados y las frecuencias correspondientes, el número de posiciones sinónimas (S) y no sinónimas (N) respecto a la secuencia máster. Para cada posición mutada se determinó si la mutación era codón sinónimo (mS) o no

sinónimo (mN) respecto a la master, teniendo en cuenta las frecuencias de cada haplotipo. La relación de posiciones no sinónimas a sinónimas, se compara con la relación mN/mS, y de la misma forma se comparan las relaciones mN/N y mS/S, como aproximación a la relación dN/dS. La relación que existe entre ambos estimadores permite determinar si las secuencias se encuentran bajo el modelo neutral, donde se esperaría un valor $(mN/N) / (mS/S)$ próximo a la unidad, o se desvían del mismo. Este valor puede ayudar a distinguir los crónicos de los agudos.

Sliding window

Aparte del análisis posición por posición es interesante evaluar la secuencia completa por pequeños segmentos, llamados ventanas, que se desplazan por pasos, hasta cubrir toda la región ("sliding window"). Como en el análisis posición por posición los reads se conectan de forma óptima, teniendo en cuenta el número de reads que pertenece a cada amplicón, los reads se recortan de acuerdo con esta conexión óptima, y las ventanas se desliza desde 5' a 3' en pasos de pocos nucleótidos. En el sliding window se representó el parámetro de diversidad nucleotídica (P_i).

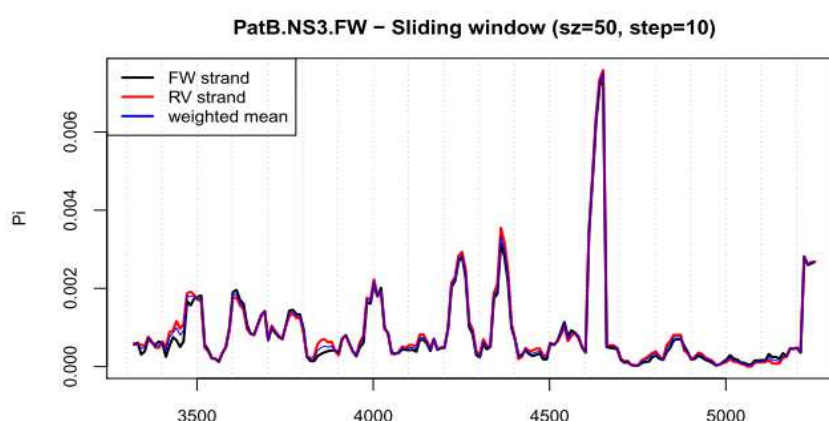


Figura 29. Representación gráfica de sliding windows de nucleótidos donde en el eje de abscisas se representó la P_i (diversidad nucleotídica) y en el eje de ordenadas se representaron las

Se desarrollaron sliding windows a nivel de nucleótidos (el tamaño de la ventana es 50 con pasos de 10) y de aminoácidos (el tamaño de la ventana es de 12 con pasos de 3).

COMPARACIÓN DE 2 POBLACIONES: A vs B

Comparación a nivel de nucleótidos

Se realizó una comparación mediante gráficos de barras de los sitios polimórficos de 2 poblaciones virales. La población en segundo plano se representó en gris y la población en primer plano, se representó cada nucleótido en un color (en el eje de abscisas se representó las posiciones nucleotídicas del gen NS3 y en el eje de ordenadas se representó el porcentaje de mutación de cada nucleótido dentro de la población).

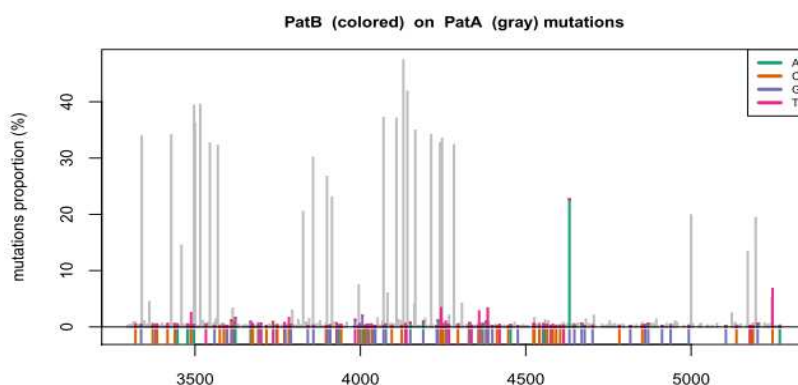


Figura 30. Gráfico de barras con comparación de frecuencias relativas de mutaciones de nucleótidos de 2 pacientes (un paciente representado en gris y otro paciente en colores diferentes).

Comparación a nivel de aminoácidos

Cada aminoácido se representó de acuerdo a sus propiedades químicas, mediante 5 colores.

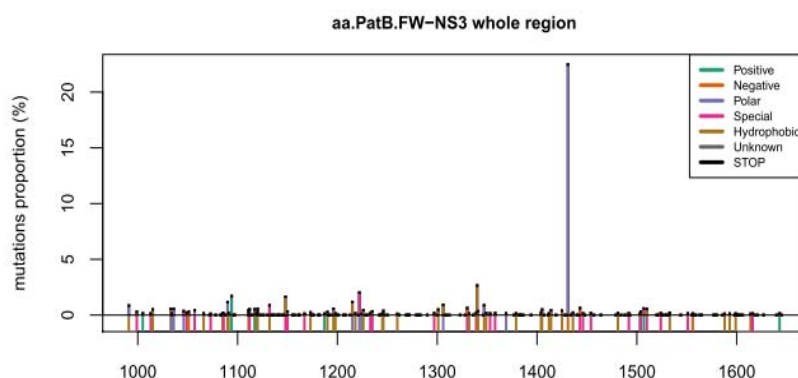


Figura 30. Gráfico de barras con comparación de frecuencias relativas de mutaciones de aminoácidos de 2 pacientes.

Identidad poblacional. Coeficiente de Bhattacharyya

Una población viral puede caracterizarse por el conjunto de pares (h_i, f_i) , $i:1..n$, donde h_i es el haplotipo i -ésimo, y f_i su frecuencia relativa en la población. Dadas 2 poblaciones se utiliza el coeficiente de Bhattachayya como medida de identidad poblacional, definido como:

$$BC = \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} \delta(h_{Ai}, h_{Bj}) \sqrt{f_{Ai} f_{Bj}}$$

Donde δ es delta Kronecker, con valor 1 cuando los dos haplotipos son iguales, o cero, en caso contrario, y las sumas se extienden a todos los haplotipos de ambas poblaciones. Este coeficiente da una aproximación del solapamiento de la población, de tal manera que cuando es 1, las poblaciones son idénticas, y cuando es cero, no existe ningún haplotipo idéntico.

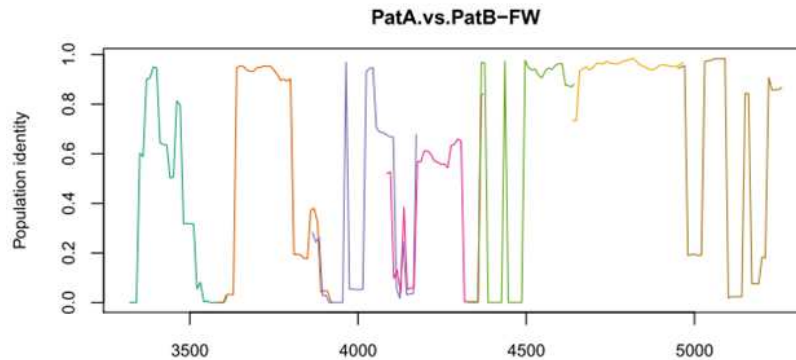


Figura 31. Representación gráfica de similitud de 2 poblaciones virales de 2 pacientes (ejemplo de A y B).

Distancia genética

Supongamos que hay m haplotipos diferentes de una región particular del ADN de las poblaciones X e Y , donde x_i e y_i son las frecuencias del haplotipo i -ésimo en la muestra para las poblaciones X e Y , respectivamente. El número promedio de sustituciones de nucleótidos para un par de haplotipos escogido de forma aleatoria (d_x) en la población X puede ser estimada como:

$$\hat{d}_x = \frac{n_X}{n_X - 1} \sum_{i,j} \hat{x}_i \hat{x}_j d_{ij}$$

donde n_x es el número de secuencias de la muestra y d_{ij} es el número de sustituciones de nucleótidos por sitio entre los haplotipos i -ésimo y j -ésimo. Cuando todas las secuencias de ADN son diferentes, $x_i=1/n_x$. El número medio de sustituciones de nucleótidos para Y puede ser estimado de la misma manera.

Por otra parte, el número medio de sustituciones de nucleótidos entre los haplotipos de ADN de X e Y puede ser estimado por:

$$\hat{d}_{XY} = \sum_{ij} x_i x_j d_{ij}$$

donde d_{ij} es las sustituciones de nucleótidos entre el haplotipo i -ésimo de X y el haplotipo j -ésimo de Y. El número neto de sustituciones de nucleótidos entre las 2 poblaciones (d_A) es entonces estimado por:

$$\hat{d}_A = \hat{d}_{XY} - (\hat{d}_X + \hat{d}_Y)/2$$

BIBLIOGRAFÍA

9 BIBLIOGRAFÍA

- (1) Penin F, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM. Structural biology of hepatitis C virus. *Hepatology* 2004 January;39(1):5-19.
- (2) Grakoui A, Wychowski C, Lin C, Feinstone SM, Rice CM. Expression and identification of hepatitis C virus polyprotein cleavage products. *J Virol* 1993 March;67(3):1385-95.
- (3) Kato N. Molecular virology of hepatitis C virus. *Acta Med Okayama* 2001 June;55(3):133-59.
- (4) Rosenberg S. Recent advances in the molecular biology of hepatitis C virus. *J Mol Biol* 2001 October 26;313(3):451-64.
- (5) Montserret R, Saint N, Vanbelle C, Salvay AG, Simorre JP, Ebel C, Sapay N, Renisio JG, Bockmann A, Steinmann E, Pietschmann T, Dubuisson J, Chipot C, Penin F. NMR structure and ion channel activity of the p7 protein from hepatitis C virus. *J Biol Chem* 2010 October 8;285(41):31446-61.
- (6) Steinmann E, Penin F, Kallis S, Patel AH, Bartenschlager R, Pietschmann T. Hepatitis C virus p7 protein is crucial for assembly and release of infectious virions. *PLoS Pathog* 2007 July;3(7):e103.
- (7) Jirasko V, Montserret R, Lee JY, Gouttenoire J, Moradpour D, Penin F, Bartenschlager R. Structural and functional studies of nonstructural protein 2 of the hepatitis C virus reveal its key role as organizer of virion assembly. *PLoS Pathog* 2010;6(12):e1001233.
- (8) Gouttenoire J, Penin F, Moradpour D. Hepatitis C virus nonstructural protein 4B: a journey into unexplored territory. *Rev Med Virol* 2010 March;20(2):117-29.
- (9) Hanouille X, Badillo A, Verdegem D, Penin F, Lippens G. The domain 2 of the HCV NS5A protein is intrinsically unstructured. *Protein Pept Lett* 2010 August;17(8):1012-8.
- (10) Verdegem D, Badillo A, Wieruszeski JM, Landrieu I, Leroy A, Bartenschlager R, Penin F, Lippens G, Hanouille X. Domain 3 of NS5A protein from the hepatitis C virus has intrinsic alpha-helical propensity and is a substrate of cyclophilin A. *J Biol Chem* 2011 June 10;286(23):20441-54.

- (11) Moradpour D, Brass V, Bieck E, Friebe P, Gosert R, Blum HE, Bartenschlager R, Penin F, Lohmann V. Membrane association of the RNA-dependent RNA polymerase is essential for hepatitis C virus RNA replication. *J Virol* 2004 December;78(23):13278-84.
- (12) Argentini C, Genovese D, Dettori S, Rapicetta M. HCV genetic variability: from quasispecies evolution to genotype classification. *Future Microbiol* 2009 April;4(3):359-73.
- (13) Ramia S, Eid-Fares J. Distribution of hepatitis C virus genotypes in the Middle East. *Int J Infect Dis* 2006 July;10(4):272-7.
- (14) Cornberg M, Manns MP. [Current Guidelines for treatment of hepatitis B. Correct diagnosis is the prerequisite.]. *Pharm Unserer Zeit* 2011 January;40(1):45-51.
- (15) Esteban JI, Sauleda S, Quer J. The changing epidemiology of hepatitis C virus infection in Europe. *J Hepatol* 2008 January;48(1):148-62.
- (16) Re V, Contigiani M, Yoshida CF, Lampe E. Identification of hepatitis C virus subtype 2c by sequencing analysis in patients from Cordoba, Argentina. *Mem Inst Oswaldo Cruz* 2007 December;102(8):995-8.
- (17) Sievert W, Altraif I, Razavi HA, Abdo A, Ahmed EA, Alomair A, Amarapurkar D, Chen CH, Dou X, El KH, Elshazly M, Esmat G, Guan R, Han KH, Koike K, Largen A, McCaughan G, Mogawer S, Monis A, Nawaz A, Piratvisuth T, Sanai FM, Sharara AI, Sibbel S, Sood A et al. A systematic review of hepatitis C virus epidemiology in Asia, Australia and Egypt. *Liver Int* 2011 July;31 Suppl 2:61-80.
- (18) Yu ML, Chuang WL. Treatment of chronic hepatitis C in Asia: when East meets West. *J Gastroenterol Hepatol* 2009 March;24(3):336-45.
- (19) Steinhauer DA, Domingo E, Holland JJ. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 1992 December 15;122(2):281-8.
- (20) Friedberg EC. A brief history of the DNA repair field. *Cell Res* 2008 January;18(1):3-7.
- (21) Batschelet E, Domingo E, Weissmann C. The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. *Gene* 1976;1(1):27-32.
- (22) Drake JW, Holland JJ. Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A* 1999 November 23;96(24):13910-3.

- (23) Okamoto H, Kojima M, Okada S, Yoshizawa H, Iizuka H, Tanaka T, Muchmore EE, Peterson DA, Ito Y, Mishiro S. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. *Virology* 1992;190:894-9.
- (24) Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, Layden TJ, Perelson AS. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *science* 1998 October 2;282(5386):103-7.
- (25) Martell M, Esteban JI, Quer J, Genesca J, Weiner A, Esteban R, Guardia J, Gomez J. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J Virol* 1992;66(5):3225-9.
- (26) Domingo E, Holland JJ. High error rates, population equilibrium, and evolution of RNA replication systems. In: Domingo E, Holland JJ, editors. *RNA Genetics*. Boca Raton, Florida: CRC Press; 1988. p. 3-36.
- (27) Eigen M, McCaskill J, Schuster P. The molecular quasi-species. In: Prigogine and S.S.Rice., editor. *Advances in Chemical Physics*. Vol. LXXV. New York: Wiley.; 1989. p. 149-263.
- (28) Domingo E, Escarmis C, Sevilla N, Moya A, Elena SF, Quer J, Novella IS, Holland JJ. Basic concepts in RNA virus evolution. *FASEB J* 1996 June;10(8):859-64.
- (29) Domingo E, Flavell RA, Weissmann C. In vitro site-directed mutagenesis: generation and properties of an infectious extracistronic mutant of bacteriophage Qbeta. *Gene* 1976;1(1):3-25.
- (30) Domingo E, Martin V, Perales C, Grande-Perez A, Garcia-Arriaza J, Arias A. Viruses as quasispecies: biological implications. *Curr Top Microbiol Immunol* 2006;299:51-82.
- (31) Fishman SL, Branch AD. The quasispecies nature and biological implications of the hepatitis C virus. *Infect Genet Evol* 2009 December;9(6):1158-67.
- (32) Holland JJ, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S. Rapid evolution of RNA genomes. *Science* 1982;215(4540):1577-85.
- (33) Holland JJ. Transitions in understanding of RNA viruses: a historical perspective. *Curr Top Microbiol Immunol* 2006;299:371-401.
- (34) Novella IS. Contributions of vesicular stomatitis virus to the understanding of RNA virus evolution. *Curr Opin Microbiol* 2003 August;6(4):399-405.

- (35) Biebricher CK, Eigen M. What is a quasispecies? *Curr Top Microbiol Immunol* 2006;299:1-31.
- (36) Duarte EA, Novella IS, Weaver SC, Domingo E, Wain-Hobson S, Clarke DK, Moya A, Elena SF, De La Torre JC, Holland JJ. RNA virus quasispecies: significance for viral disease and epidemiology. *Infect Agents Dis* 1994 August;3(4):201-14.
- (37) Holland JJ, De La Torre JC, Steinhauer DA. RNA virus populations as quasispecies. *Curr Top Microbiol Immunol* 1992;176:1-20.
- (38) Steinhauer DA, Holland JJ. Rapid evolution of RNA viruses. *Annu Rev Microbiol* 1987;41:409-33.
- (39) Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 2006 January 19;439(7074):344-8.
- (40) Domingo E. *Virus en evolución*. Madrid: Eudema S.A.; 1994.
- (41) Mikkelsen JG, Pedersen FS. Genetic reassortment and patch repair by recombination in retroviruses. *J Biomed Sci* 2000 March;7(2):77-99.
- (42) Kalinina O, Norder H, Mukomolov S, Magnus LO. A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J Virol* 2002 April;76(8):4034-43.
- (43) Gonzalez-Candelas F, Lopez-Labrador FX, Bracho MA. Recombination in hepatitis C virus. *Viruses* 2011 October;3(10):2006-24.
- (44) De La Torre JC, Holland JJ. RNA virus quasispecies populations can suppress vastly superior mutant progeny. *J Virol* 1990 December;64(12):6278-81.
- (45) Holland JJ, De La Torre JC, Clarke DK, Duarte E. Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. *J Virol* 1991;65:2960-7.
- (46) Martinez MA, Carrillo C, Gonzalez-Candelas F, Moya A, Domingo E, Sobrino F. Fitness alteration of foot-and-mouth disease virus mutants: measurement of adaptability of viral quasispecies. *J Virol* 1991 July;65(7):3954-7.
- (47) Domingo E, Holland JJ. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* 1997;51:151-78.

- (48) Novella IS, Elena SF, Moya A, Domingo E, Holland JJ. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. *J Virol* 1995 May;69(5):2869-72.
- (49) Escarmis C, Davila M, Domingo E. Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J Mol Biol* 1999 January 15;285(2):495-505.
- (50) Clarke DK, Duarte EA, Moya A, Elena SF, Domingo E, Holland J. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *J Virol* 1993 January;67(1):222-8.
- (51) Duarte E, Clarke D, Moya A, Domingo E, Holland J. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc Natl Acad Sci U S A* 1992 July 1;89(13):6015-9.
- (52) Escarmis C, Davila M, Charpentier N, Bracho A, Moya A, Domingo E. Genetic lesions associated with Muller's ratchet in an RNA virus. *J Mol Biol* 1996 November 29;264(2):255-67.
- (53) Alter HJ, Conry-Cantilena C, Melpolder J, Tan D, Van Raden M, Herion D, Lau D, Hoofnagle JH. Hepatitis C in asymptomatic blood donors. *Hepatology* 1997 September;26(3 Suppl 1):29S-33S.
- (54) Llaure AS, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* 2010;6(7):e1001005.
- (55) Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, Melpolder JC, Strazzer A, Chien DY, Munoz SJ, Balestrieri A, Purcell RH, Alter HJ. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *science* 2000 April 14;288(5464):339-44.
- (56) Grande Gimenez MC, El Far F, Barsanti WS, Servolo Medeiros EA. Cut and puncture accidents involving health care workers exposed to biological materials. *Braz J Infect Dis* 2001 October;5(5):235-42.
- (57) Puro V, Petrosillo N, Ippolito G, Aloisi MS, Boumis E, Rava L. Occupational hepatitis C virus infection in Italian health care workers. Italian Study Group on Occupational Risk of Bloodborne Infections. *Am J Public Health* 1995 September;85(9):1272-5.
- (58) Quer J, Murillo P, Esteban JI, Martell M, Esteban R, Guardia J. Sexual transmission of hepatitis C virus from a patient with chronic disease to his sex partner after removal of an intrauterine device. *Sex Transm Dis* 2003 May;30(5):470-1.
- (59) Roberts EA, Yeung L. Maternal-infant transmission of hepatitis C virus infection. *Hepatology* 2002 November;36(5 Suppl 1):S106-S113.

- (60) Ross RS, Viazov S, Thormahlen M, Bartz L, Tamm J, Rautenberg P, Roggendorf M, Deister A. Risk of hepatitis C virus transmission from an infected gynecologist to patients: results of a 7-year retrospective investigation. *Arch Intern Med* 2002 April 8;162(7):805-10.
- (61) Bronowicki JP, Vetter D, Uhl G, Hudziak H, Uhrlacher A, Vetter JM, Doffoel M. Lymphocyte reactivity to hepatitis C virus (HCV) antigens shows evidence for exposure to HCV in HCV-seronegative spouses of HCV-infected patients. *J Infect Dis* 1997 August;176(2):518-22.
- (62) Chao L. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 1990 November 29;348(6300):454-5.
- (63) Domingo E, Escarmis C, Sevilla N, Moya A, Elena SF, Quer J, Novella IS, Holland JJ. Basic concepts in RNA virus evolution. *Faseb Journal* 1996;10(8):859-64.
- (64) Duarte EA, Clarke DK, Moya A, Elena SF, Domingo E, Holland J. Many-trillionfold amplification of single RNA virus particles fails to overcome the Muller's ratchet effect. *J Virol* 1993 June;67(6):3620-3.
- (65) Quer J, Murillo P, Esteban JI, Martell M, Esteban R, Guardia J. Sexual transmission of hepatitis C virus from a patient with chronic disease to his sex partner after removal of an intrauterine device. *Sexually Transmitted Diseases* 2003;30(5):470-1.
- (66) Domingo E, Gomez J. Quasispecies and its impact on viral hepatitis. *Virus Res* 2007 March 7;127(2):131-50.
- (67) Cooper S, Erickson AL, Adams EJ, Kansopon J, Weiner AJ, Chien DY, Houghton M, Parham P, Walker CM. Analysis of a successful immune response against hepatitis C virus. *Immunity* 1999 April;10(4):439-49.
- (68) Erickson AL, Kimura Y, Igarashi S, Eichelberger J, Houghton M, Sidney J, McKinney D, Sette A, Hughes AL, Walker CM. The outcome of hepatitis C virus infection is predicted by escape mutations in epitopes targeted by cytotoxic T lymphocytes. *Immunity* 2001 December;15(6):883-95.
- (69) Kawada M, Igarashi H, Takeda A, Tsukamoto T, Yamamoto H, Dohki S, Takiguchi M, Matano T. Involvement of multiple epitope-specific cytotoxic T-lymphocyte responses in vaccine-based control of simian immunodeficiency virus replication in rhesus macaques. *J Virol* 2006 February;80(4):1949-58.
- (70) Quinones-Mateu ME, Arts EJ. Virus fitness: concept, quantification, and application to HIV population dynamics. *Curr Top Microbiol Immunol* 2006;299:83-140.

- (71) Yi M, Tong X, Skelton A, Chase R, Chen T, Prongay A, Bogen SL, Saksena AK, Njoroge FG, Veselenak RL, Pyles RB, Bourne N, Malcolm BA, Lemon SM. Mutations conferring resistance to SCH6, a novel hepatitis C virus NS3/4A protease inhibitor. Reduced RNA replication fitness and partial rescue by second-site mutations. *J Biol Chem* 2006 March 24;281(12):8205-15.
- (72) Ruiz-Jarabo CM, Arias A, Baranowski E, Escarmís C, Domingo E. Memory in viral quasispecies. *J Virol* 2000 April;74(8):3543-7.
- (73) Susser S, Vermehren J, Forestier N, Welker MW, Grigorian N, Fuller C, Perner D, Zeuzem S, Sarrazin C. Analysis of long-term persistence of resistance mutations within the hepatitis C virus NS3 protease after treatment with telaprevir or boceprevir. *J Clin Virol* 2011 December;52(4):321-7.
- (74) Thio CL. Host genetic factors and antiviral immune responses to hepatitis C virus. *Clin Liver Dis* 2008 August;12(3):713-26, xi.
- (75) Asselah T, Bieche I, Paradis V, Bedossa P, Vidaud M, Marcellin P. Genetics, genomics, and proteomics: implications for the diagnosis and the treatment of chronic hepatitis C. *Semin Liver Dis* 2007 February;27(1):13-27.
- (76) Massard J, Ratziu V, Thabut D, Moussalli J, Lebray P, Benhamou Y, Poynard T. Natural history and predictors of disease severity in chronic hepatitis C. *J Hepatol* 2006;44(1 Suppl):S19-S24.
- (77) Rauch A, Kutalik Z, Descombes P, Cai T, Di IJ, Mueller T, Bochud M, Battegay M, Bernasconi E, Borovicka J, Colombo S, Cerny A, Dufour JF, Furrer H, Gunthard HF, Heim M, Hirschel B, Malinverni R, Moradpour D, Mullhaupt B, Witteck A, Beckmann JS, Berg T, Bergmann S, Negro F et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* 2010 April;138(4):1338-45, 1345.
- (78) Afdhal NH, McHutchison JG, Zeuzem S, Mangia A, Pawlotsky JM, Murray JS, Shianna KV, Tanaka Y, Thomas DL, Booth DR, Goldstein DB. Hepatitis C pharmacogenetics: state of the art in 2010. *Hepatology* 2011 January;53(1):336-45.
- (79) Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009 September 17;461(7262):399-401.
- (80) Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, Kidd J, Kidd K, Khakoo SI, Alexander G, Goedert JJ, Kirk GD, Donfield SM, Rosen HR, Tobler LH, Busch MP, McHutchison JG, Goldstein DB, Carrington M.

Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* 2009 October 8;461(7265):798-801.

- (81) Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K, Masaki N et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009 October;41(10):1105-9.
- (82) Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E, Riordan S, Sheridan D, Smedile A, Fragomeli V, Muller T, Bahlo M, Stewart GJ, Booth DR, George J. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009 October;41(10):1100-4.
- (83) Bigger CB, Brasky KM, Lanford RE. DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J Virol* 2001 August;75(15):7059-66.
- (84) Su AI, Pezacki JP, Wodicka L, Brideau AD, Supekova L, Thimme R, Wieland S, Bukh J, Purcell RH, Schultz PG, Chisari FV. Genomic analysis of the host response to hepatitis C virus infection. *Proc Natl Acad Sci U S A* 2002 November 26;99(24):15669-74.
- (85) Takaki A, Wiese M, Maertens G, Depla E, Seifert U, Liebetrau A, Miller JL, Manns MP, Rehermann B. Cellular immune responses persist and humoral responses decrease two decades after recovery from a single-source outbreak of hepatitis C. *Nat Med* 2000 May;6(5):578-82.
- (86) Koziel MJ, Wong DK, Dudley D, Houghton M, Walker BD. Hepatitis C virus-specific cytolytic T lymphocyte and T helper cell responses in seronegative persons. *J Infect Dis* 1997 October;176(4):859-66.
- (87) Day CL, Lauer GM, Robbins GK, McGovern B, Wurcel AG, Gandhi RT, Chung RT, Walker BD. Broad specificity of virus-specific CD4+ T-helper-cell responses in resolved hepatitis C virus infection. *J Virol* 2002 December;76(24):12584-95.
- (88) Diepolder HM, Zachoval R, Hoffmann RM, Wierenga EA, Santantonio T, Jung MC, Eichenlaub D, Pape GR. Possible mechanism involving T-lymphocyte response to non-structural protein 3 in viral clearance in acute hepatitis C virus infection. *Lancet* 1995 October 14;346(8981):1006-7.
- (89) Diepolder HM, Gerlach JT, Zachoval R, Hoffmann RM, Jung MC, Wierenga EA, Scholz S, Santantonio T, Houghton M, Southwood S, Sette A, Pape GR. Immunodominant CD4+ T-cell epitope within nonstructural protein 3 in acute hepatitis C virus infection. *J Virol* 1997 August;71(8):6011-9.

- (90) Gerlach JT, Diepolder HM, Jung MC, Gruener NH, Schraut WW, Zachoval R, Hoffmann R, Schirren CA, Santantonio T, Pape GR. Recurrence of hepatitis C virus after loss of virus-specific CD4(+) T-cell response in acute hepatitis C. *Gastroenterology* 1999 October;117(4):933-41.
- (91) Gruner NH, Gerlach JT, Jung MC, Diepolder HM, Schirren CA, Schraut WW. Association of hepatitis C virus-specific CD8+ T cells with viral clearance in acute hepatitis C. *J Infect Dis* 2000;181:1528-36.
- (92) Lechner F, Wong DK, Dunbar PR, Chapman R, Chung RT, Dohrenwend P, Robbins G, Phillips R, Klenerman P, Walker BD. Analysis of successful immune responses in persons infected with hepatitis C virus. *J Exp Med* 2000 May 1;191(9):1499-512.
- (93) Missale G, Bertoni R, Lamonaca V, Valli A, Massari M, Mori C, Rumi MG, Houghton M, Fiaccadori F, Ferrari C. Different clinical behaviors of acute hepatitis C virus infection are associated with different vigor of the anti-viral cell-mediated immune response. *J Clin Invest* 1996 August 1;98(3):706-14.
- (94) Urbani S, Amadei B, Fisicaro P, Tola D, Orlandini A, Sacchelli L, Mori C, Missale G, Ferrari C. Outcome of acute hepatitis C is related to virus-specific CD4 function and maturation of antiviral memory CD8 responses. *Hepatology* 2006 July;44(1):126-39.
- (95) Lamonaca V, Missale G, Urbani S, Pilli M, Boni C, Mori C, Sette A, Massari M, Southwood S, Bertoni R, Valli A, Fiaccadori F, Ferrari C. Conserved hepatitis C virus sequences are highly immunogenic for CD4(+) T cells: implications for vaccine development. *Hepatology* 1999 October;30(4):1088-98.
- (96) Alric L, Fort M, Izopet J, Vinel JP, Charlet JP, Selves J, Puel J, Pascal JP, Duffaut M, Abbai M. Genes of the major histocompatibility complex class II influence the outcome of hepatitis C virus infection. *Gastroenterology* 1997 November;113(5):1675-81.
- (97) McKiernan SM, Hagan R, Curry M, McDonald GS, Nolan N, Crowley J, Hegarty J, Lawlor E, Kelleher D. The MHC is a major determinant of viral status, but not fibrotic stage, in individuals infected with hepatitis C. *Gastroenterology* 2000 June;118(6):1124-30.
- (98) Minton EJ, Smillie D, Neal KR, Irving WL, Underwood JC, James V. Association between MHC class II alleles and clearance of circulating hepatitis C virus. Members of the Trent Hepatitis C Virus Study Group. *J Infect Dis* 1998 July;178(1):39-44.

- (99) Thio CL, Gao X, Goedert JJ, Vlahov D, Nelson KE, Hilgartner MW, O'Brien SJ, Karacki P, Astemborski J, Carrington M, Thomas DL. HLA-Cw*04 and hepatitis C virus persistence. *J Virol* 2002 May;76(10):4792-7.
- (100) Thursz M, Yallop R, Goldin R, Trepo C, Thomas HC. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. The HENCORE group. Hepatitis C European Network for Cooperative Research. *Lancet* 1999 December 18;354(9196):2119-24.
- (101) Bengsch B, Thimme R, Blum HE. Role of host genetic factors in the outcome of hepatitis C virus infection. *Viruses* 2009 September;1(2):104-25.
- (102) Wang JH, Layden TJ, Eckels DD. Modulation of the peripheral T-Cell response by CD4 mutants of hepatitis C virus: transition from a Th1 to a Th2 response. *Hum Immunol* 2003 July;64(7):662-73.
- (103) Bennett S. Solexa Ltd. *Pharmacogenomics* 2004 June;5(4):433-8.
- (104) Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 2011 July 21;475(7356):348-52.
- (105) Chevaliez S, Rodriguez C, Pawlotsky JM. New virologic tools for management of chronic hepatitis B and C. *Gastroenterology* 2012 May;142(6):1303-13.
- (106) Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci U S A* 2009 May 12;106(19):7702-7.
- (107) Leruez-Ville M, Kunstmann JM, De Almeida M, Rouzioux C, Chaix ML. Detection of hepatitis C virus in the semen of infected men. *Lancet* 2000 July 1;356(9223):42-3.
- (108) Bartels DJ, Zhou Y, Zhang EZ, Marcial M, Byrn RA, Pfeiffer T, Tigges AM, Adiwijaya BS, Lin C, Kwong AD, Kieffer TL. Natural prevalence of hepatitis C virus variants with decreased sensitivity to NS3.4A protease inhibitors in treatment-naïve subjects. *J Infect Dis* 2008 September 15;198(6):800-7.
- (109) Colson P, Brouk N, Lembo F, Castellani P, Tamalet C, Gerolami R. Natural presence of substitution R155K within hepatitis C virus NS3 protease from a treatment-naïve chronically infected patient. *Hepatology* 2008 February;47(2):766-7.

- (110) Kim AY, Timm J, Nolan BE, Reyor LL, Kane K, Berical AC, Zachary KC, Lauer GM, Kuntzen T, Allen TM. Temporal dynamics of a predominant protease inhibitor-resistance mutation in a treatment-naïve, hepatitis C virus-infected individual. *J Infect Dis* 2009 March 1;199(5):737-41.
- (111) Kuntzen T, Timm J, Berical A, Lennon N, Berlin AM, Young SK, Lee B, Heckerman D, Carlson J, Reyor LL, Kleyman M, McMahon CM, Birch C, Schulze zur WJ, Ledlie T, Koehrsen M, Kodira C, Roberts AD, Lauer GM, Rosen HR, Bihl F, Cerny A, Spengler U, Liu Z, Kim AY et al. Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naïve patients. *Hepatology* 2008 December;48(6):1769-78.
- (112) McKiernan SM, Hagan R, Curry M, McDonald GS, Kelly A, Nolan N, Walsh A, Hegarty J, Lawlor E, Kelleher D. Distinct MHC class I and II alleles are associated with hepatitis C viral clearance, originating from a single source. *Hepatology* 2004 July;40(1):108-14.
- (113) Wertheimer AM, Miner C, Lewinsohn DM, Sasaki AW, Kaufman E, Rosen HR. Novel CD4+ and CD8+ T-cell determinants within the NS3 protein in subjects with spontaneously resolved HCV infection. *Hepatology* 2003 March;37(3):577-89.
- (114) He Y, King MS, Kempf DJ, Lu L, Lim HB, Krishnan P, Kati W, Middleton T, Molla A. Relative replication capacity and selective advantage profiles of protease inhibitor-resistant hepatitis C virus (HCV) NS3 protease mutants in the HCV genotype 1b replicon system. *Antimicrob Agents Chemother* 2008 March;52(3):1101-10.
- (115) Kieffer TL, Sarrazin C, Miller JS, Welker MW, Forestier N, Reesink HW, Kwong AD, Zeuzem S. Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis C virus replication in patients. *Hepatology* 2007 September;46(3):631-9.
- (116) Sarrazin C, Kieffer TL, Bartels D, Hanzelka B, Muh U, Welker M, Wincheringer D, Zhou Y, Chu HM, Lin C, Weegink C, Reesink H, Zeuzem S, Kwong AD. Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology* 2007 May;132(5):1767-77.
- (117) Welsch C, Domingues FS, Susser S, Antes I, Hartmann C, Mayr G, Schlicker A, Sarrazin C, Albrecht M, Zeuzem S, Lengauer T. Molecular basis of telaprevir resistance due to V36 and T54 mutations in the NS3-4A protease of the hepatitis C virus. *Genome Biol* 2008;9(1):R16.
- (118) Tong X, Bogen S, Chase R, Girijavallabhan V, Guo Z, Njoroge FG, Prongay A, Saksena A, Skelton A, Xia E, Ralston R. Characterization of resistance

- mutations against HCV ketoamide protease inhibitors. *Antiviral Res* 2008 March;77(3):177-85.
- (119) Sarrazin C, Rouzier R, Wagner F, Forestier N, Larrey D, Gupta SK, Hussain M, Shah A, Cutler D, Zhang J, Zeuzem S. SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon alpha-2b for genotype 1 nonresponders. *Gastroenterology* 2007 April;132(4):1270-8.
- (120) Qiu P, Sanfiorenzo V, Curry S, Guo Z, Liu S, Skelton A, Xia E, Cullen C, Ralston R, Greene J, Tong X. Identification of HCV protease inhibitor resistance mutations by selection pressure-based method. *Nucleic Acids Res* 2009 June;37(10):e74.
- (121) Susser S, Welsch C, Wang Y, Zettler M, Domingues FS, Karey U, Hughes E, Ralston R, Tong X, Herrmann E, Zeuzem S, Sarrazin C. Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients. *Hepatology* 2009 December;50(6):1709-18.
- (122) Aparicio E, Franco S, Parera M, Andres C, Tural C, Clotet B, Martinez MA. Complexity and catalytic efficiency of hepatitis C virus (HCV) NS3 and NS4A protease quasispecies influence responsiveness to treatment with pegylated interferon plus ribavirin in HCV/HIV-coinfected patients. *J Virol* 2011 June;85(12):5961-9.
- (123) Puig-Basagoiti, F, Sáiz JC, Forns X, Ampurdanès S, Giménez-Barcons M, Franco, S, Sánchez-Fueyo, A, Costa J, Sanchez-Tapias JM, Rodés J. 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* 2001;65(1):35-44.
- (124) Feliu A, Gay E, Garcia-Retortillo M, Saiz JC, Forns X. Evolution of hepatitis C virus quasispecies immediately following liver transplantation. *Liver Transpl* 2004 September;10(9):1131-9.
- (125) Garcia-Retortillo M, Forns X, Feliu A, Moitinho E, Costa J, Navasa M, Rimola A, Rodes J. Hepatitis C virus kinetics during and immediately after liver transplantation. *Hepatology* 2002 March;35(3):680-7.
- (126) Hughes MG, Jr., Rudy CK, Chong TW, Smith RL, Evans HL, Iezzoni JC, Sawyer RG, Pruett TL. E2 quasispecies specificity of hepatitis C virus association with allografts immediately after liver transplantation. *Liver Transpl* 2004 February;10(2):208-16.
- (127) Gerlach JT, Diepolder HM, Zachoval R, Gruener NH, Jung MC, Ulsenheimer A, Schraut WW, Schirren CA, Waechtler M, Backmund M, Pape GR. Acute hepatitis C: high rate of both spontaneous and treatment-induced viral clearance. *Gastroenterology* 2003 July;125(1):80-8.

- (128) Grakoui A, Shoukry NH, Woollard DJ, Han JH, Hanson HL, Ghayeb J, Murthy KK, Rice CM, Walker CM. HCV persistence and immune evasion in the absence of memory T cell help. *science* 2003 October 24;302(5645):659-62.
- (129) Puig M, Mihalik K, Tilton JC, Williams O, Merchlinsky M, Connors M, Feinstone SM, Major ME. CD4+ immune escape and subsequent T-cell failure following chimpanzee immunization against hepatitis C virus. *Hepatology* 2006 September;44(3):736-45.
- (130) Sadegh-Nasseri S, Dalai SK, Korb Ferris LC, Mirshahidi S. Suboptimal engagement of the T-cell receptor by a variety of peptide-MHC ligands triggers T-cell anergy. *Immunology* 2010 January;129(1):1-7.
- (131) Brooks DG, Teyton L, Oldstone MB, McGavern DB. Intrinsic functional dysregulation of CD4 T cells occurs rapidly following persistent viral infection. *J Virol* 2005 August;79(16):10514-27.
- (132) Bes M, Sauleda S, Casamitjana N, Piron M, Campos-Varela I, Quer J, Cubero M, Puig L, Guardia J, Esteban JI. Reversal of nonstructural protein 3-specific CD4(+) T cell dysfunction in patients with persistent hepatitis C virus infection. *J Viral Hepat* 2012 April;19(4):283-94.
- (133) Farci P. New insights into the HCV quasispecies and compartmentalization. *Semin Liver Dis* 2011 November;31(4):356-74.
- (134) Pawlotsky JM. Hepatitis C virus population dynamics during infection. *Curr Top Microbiol Immunol* 2006;299:261-84.
- (135) Weiner A, Erickson AL, Kansopon J, Crawford K, Muchmore E, Hughes AL, Houghton M, Walker CM. Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escape variant. *Proc Natl Acad Sci U S A* 1995;92(7):2755-9.
- (136) Thimme R, Bukh J, Spangenberg HC, Wieland S, Pemberton J, Steiger C, Govindarajan S, Purcell RH, Chisari FV. Viral and immunological determinants of hepatitis C virus clearance, persistence, and disease. *Proc Natl Acad Sci U S A* 2002 November 26;99(24):15661-8.
- (137) Wagner A. Gene duplications, robustness and evolutionary innovations. *Bioessays* 2008 April;30(4):367-73.
- (138) Quer J, Esteban JI, Sanchez JM, Otero T, Rius C, Coll M, Cubero M, Moreno G, Gonzalez A, Vaque J, Esteban R, Campins M, Panella H, Guardia J, Martell M. Nosocomial transmission of hepatitis C virus during contrast-enhanced computed tomography scanning. *Eur J Gastroenterol Hepatol* 2008 January;20(1):73-8.

- (139) Cline J, Braman JC, Hogrefe HH. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res* 1996 September 15;24(18):3546-51.
- (140) Andre P, Kim A, Khrapko K, Thilly WG. Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. *Genome Res* 1997 August;7(8):843-52.
- (141) Cubero M, Esteban JI, Otero T, Saucedo S, Bes M, Esteban R, Guardia J, Quer J. Naturally occurring NS3-protease-inhibitor resistant mutant A156T in the liver of an untreated chronic hepatitis C patient. *Virology* 2008 January 20;370(2):237-45.
- (142) Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. *PLoS Genet* 2008 December;4(12):e1000304.

M^a DOLORES CUBERO LEÓN