



**Universitat de les
Illes Balears**

Departament de Biologia
Laboratori de Genètica

**Interacció VHC-hoste:
Estudi genètic i clínic en pacients coinfectats
amb VHC-VIH**

Tesi doctoral presentada per Marina Matas Crespí per optar al títol universitari oficial de Doctora per la Universitat de les Illes Balears, dins del programa de doctorat en Biotecnologia, Genètica i Biologia Cel·lular

Dr. José Aurelio Castro Ocón

Dra. M. Misericòrdia Ramon Juanpere

Marina Matas Crespí

Palma, 2012

Agraïments

I pareixia que no havia d'arribar mai... i ha arribat, el moment d'escriure els agraïments, la pàgina que, pel que he sentit, costa més temps d'escriure (i, pel que veig, és cert)... esper esmentar a tothom encara que segur que em deix gent, gent que en algun moment o un altre, m'han donat ànims i m'han fet sentir bé amb el que he estat fent, així que el primer agraïment és per ells: amics, família, coneguts, companys gràcies a tots!

Començaré els agraïments, amb noms i llinatges, amb els meus dos directors de tesi, el doctor José Aurelio Castro i la doctora Misericòrdia Ramon del laboratori de Genètica de la UIB, evidentment, indispensables i sense els que aquesta tesi no hagués estat possible.

Aquesta tesi tampoc hagués estat possible sense la Dra. Antònia Picornell, que va ser qui em va "fixar", ja fa molt de temps, com a alumna col·laboradora seva, fet que va fer que comencés aquesta aventura al laboratori de Genètica de la UIB.

De manera menys oficial, però pentura més meva, Pepe, Cori i Antònia, moltíssimes gràcies per haver-me ajudat amb tot el que vos he demanat, per haver estat al meu costat i per fer-me passar uns anys tan bons com els que he passat. Em sap greu, perquè sé que a alguns de vosaltres no vos agrada, però sempre sereu els meus "jefes". En aquest punt tampoc em puc oblidar de na Bàrbara, que completa el quartet al qui sempre es pot recórrer quan es té algun problema de qualsevol tipus.

Tota aquesta investigació tampoc hagués estat possible sense l'ajuda i col·laboració de l'equip de medicina interna de l'hospital Son Llàtzer de Palma, moltes gràcies a tots. Volia agrair molt especialment a la metgessa Carmen Cifuentes per haver estat sempre disponible perquè li fes els meus "interrogatoris" sobre els aspectes clínics del treball.

Gràcies també a l'Institut Cavanilles de Biodiversitat i Biologia evolutiva (València) i al seu director en aquell moment, el Dr. Andrés Moya, per acceptar-me per fer una estada breu en el seu laboratori on la Dra. Alma Bracho em va ensenyar tot el que va poder, en una setmana, del món del virus, a més de fer-me sentir com a casa. Crec que va ser una de les setmanes més profitoses d'aquest període d'investigació. Moltíssimes gràcies, va ser un moment clau en el començament d'aquest treball.

No em puc oblidar de la resta de professors de l'àrea de Genètica, el Dr. Eduard Petitpierre i el Dr. Carlos Juan que em varen iniciar en els coneixements d'aquesta àrea i que sempre han tengut una paraula amable per jo. Al Dr. Joan Pons per les xerrades de bon dematí. Tampoc de la resta de companys que han passat pel seu laboratori: Carlos V., Mar, José, Fernando (que sap que aquí a Palma té un company tennístic i nàutic per quan vulgui!)... i d'altres laboratoris veïns que a part de compartir bons moments i menjars, mai em varen dir que no a qualsevol necessitat d'ajuda (se m'han acabat les puntes, els guants, el que fos!!).

Una de les primeres persones que em va ajudar quan vaig arribar al laboratori va ser na Paloma, encara guard els apunts que vaig prendre quan ella m'explicava les coses... de fet, no és que els guardi, és que els he fet servir fins al darrer moment! Dedicada a ella va la seqüenciació del DNA mitocondrial d'aquest treball. Moltíssimes gràcies per explicar-me les coses tan bé i amb tanta paciència.

La persona que té el mèrit més gran d'haver-me aguantat per més temps durant aquests darrers anys, és na Virginia, quantes hores hem passat una davant l'altre concentrades amb els nostres ordinadors? No ens podíem queixar, un despatx ben gran, només per nosaltres dues, quan a altres laboratoris estaven ben extrets!! Moltes gràcies per tot, per ensenyar-me, pels bons moments i pel suport en els dolents.

I la tranquil·litat que teníem es va acabar amb l'arribada de dos terratrèmols... "ses Joanes"... primer una, després l'altra, i després les dues. No sé que vos he de dir que no vos hagi dit ja... Joana Francesca, seré sincera, ningú que et coneix es creu el que dius de que ets tímida, vergonyosa... no, ningú s'ho creu. Cerques alegria i optimisme al diccionari i surt la teva foto! Això és el que transmets! Joana Maria (Zape), qui m'hauria de dir que a aquestes alçades em descobririen que tenc una ànima bessona a Sa Pobla que em dóna a conèixer coses tan bones com les banderilles!! Bé, com que això es suposa que són els agraïments d'una tesi, també vos diré que gràcies a vosaltres, he après moltíssim, que per cada: una pregunta? que em fèieu m'havia de cercar la vida (i suar) per contestar (gràcies senyor *google*). Moltes gràcies per tot a les dues!!

També volia donar les gràcies a tota la resta de personal de l'edifici; a na Carme i la seva alegria, en Pep Miquel, en Xisco, n'Aina, n'Andreu, en Macià i molts més, per la feina que fan per facilitar-nos la nostra feina i per fer-me sentir sempre a gust amb ells.

Molt especialment volia agrair a n'Emi tota la seva companyia des de que vam començar a la UIB i les nostres vides quasi paral·leles, encara que en aquest darrer tram et vares avançar! (molt merescudament). Moltes gràcies per estar allà, pels cafès i berenars, i que en poguem seguir prenent molts!

Parlant de cafès, no em puc oblidar en Toni P., el meu amic físic, com l'anomenaven pel laboratori, el meu company més assidu de cafès (un inasc??), esper poder seguir prenent molts de cafès amb tu i continuar les nostres converses multidisciplinars (a veure si aconseguix entendre alguna cosa del teu camp!). Haurem d'anar cercant bar per Palma...

Passant de cafès a xocolata, arribam a na MariaTous, (sense espais, no és un error, és com la coneixem nosaltres). Juntes a l'escola des dels dos anys, estudiant carreres i fent tesis en edificis veïnats, es pot dir que em estat sempre juntes, a estones sense veure'ns gaire, per després quedar i estar hores i hores sense aturar de xerrar (quants de quilos de patatilles ens haurem menjat??? Millor no saber-ho!!). Què en poguem menjar molts de quilos més!

Com no podia ser d'una altra manera, a les persones que més he d'agrair estar aquí són els meus pares, Maria i Toni, perquè aquesta tesi és, com els meus gens, un 50% de cada un. Moltes gràcies per tot, per animar-me a estudiar, a seguir estudiant i a acabar el que havia començat!

També gràcies a na Laura L. i a tots els amics que no puc esmentar un a un, a la pandilla de l'escola per sempre interessar-se pel meu treball.

Moltes gràcies a la resta de família, a la padrina (i també als padrins que ja no hi són), Miquel, Xisco i Maria, Maria i Xisco, Bàrbara, Selva, Marteta... i també a la família d'en Miquel, per haver-me animat sempre i cuidar-me a base de bon menjarets.

També em falta esmentar el meu germà Toniet, una de les meves motivacions era acabar l'etapa universitària abans que tu la comencassis, però... m'has guanyat!! No vall!! Esper que disfrutis d'aquesta etapa (esper que no tan llarga, o si??) tant com jo.

Finalment, només em queda donar-li les gràcies a en Miquel, el millor amic i company de vida, viatges i batalles (perquè segur que en guanyarem moltes més junts) que mai hagués pogut trobar. Moltes gràcies per aguantar les meves curolles "investigadores" i tenir tanta paciència amb jo (sí, en tens, encara que diguis que no, però la deus gastar tota amb mi!). Moltes gràcies per començar tots els meus dies amb el millor cafè del món!

Marina

0. ÍNDIX	1
1. INTRODUCCIÓ	5
1. EL VIRUS DE L'HEPATITIS C (VHC)	7
A. ESTRUCTURA DEL VHC	7
B. GENOTIPS I VARIABILITAT	10
C. PREVALENÇA I DISTRIBUCIÓ GEOGRÀFICA	11
2. PATOLOGIA DEL VHC	13
A. HISTÒRIA NATURAL DE LA INFECCIÓ	13
B. DIAGNÒSTIC	16
C. TRACTAMENT	17
3. COINFECCIÓ VHC-VIH	19
A. VIH COM A MODIFICADOR DE LA INFECCIÓ PEL VHC	19
B. PREVALENÇA I DISTRIBUCIÓ DE LA COINFECCIÓ VHC-VIH	20
4. DANY HEPÀTIC	21
A. FIBROSI	21
B. MESURES NO INVASIVES DEL DANY HEPÀTIC	23
5. VARIABILITAT GENÈTICA DELS HOSTES	23
A. IL28B (RS12979860)	24
B. CTLA4 (RS5742909 I RS231775)	26
C. LDLR (RS14158)	27
D. HFE	28
E. HAPLOGRUPS MITOCONDRIALS	29
2. OBJECTIUS	31
3. MATERIAL I MÈTODES	35
1. ESTUDI DEL GENOMA VÍRIC <i>versus</i> DANY HEPÀTIC	37
A. PACIENTS	37
B. VARIABLES	37
C. MESURA DE LA FIBROSI HEPÀTICA	38
D. SEQÜENCIACIÓ I FILOGÈNIA DE LES REGIONS <i>CORE</i> I <i>NS5B</i>	38
E. ANÀLISIS ESTADÍSTIQUES	42
2. ESTUDI DELS POLIMORFISMES EN PACIENTS <i>versus</i> RESPOSTA AL TRACTAMENT	43
A. PACIENTS I MOSTRES CONTROL	43
B. VARIABLES	44
C. DETECCIÓ DELS POLIMORFISMES	44
<i>Extracció del DNA</i>	44
<i>Detecció dels polimorfismes</i>	44
<i>RFLPs: digestió i gels d'agarosa</i>	47
<i>Seqüenciació</i>	49
D. ANÀLISIS ESTADÍSTIQUES	50

4. PUBLICACIONS	51
0. ÍNDEX DE PUBLICACIONS	53
1. MATAS M, PICORNELL A, CIFUENTES C, PAYERAS A, BASSA A, HOMAR F, LÓPEZ-LABRADOR FX, MOYA A, RAMON MM, CASTRO JA (2010). RELATING THE LIVER DAMAGE WITH HEPATITIS C VIRUS POLYMORPHISM IN CORE REGION AND HUMAN VARIABLES IN HIV-1-COINFECTED PATIENTS. INFECT GENET EVOL 10, 1252-1261.....	55
2. GENERALIZED LINEAR MODEL (GLM) FRAMEWORK FOR THE ASSOCIATION OF HOST VARIABLES AND VIRAL STRAINS WITH LIVER FIBROSIS IN HCV/HIV COINFECTED PATIENTS	81
A. ABSTRACT	82
B. INTRODUCTION.....	83
C. MATERIALS AND METHODS.....	85
D. RESULTS	88
<i>Baseline characteristics of study subjects</i>	88
<i>HCV NS5B phylogeny</i>	88
<i>GLM and other Statistical analyses</i>	89
E. DISCUSSION	90
F. CONCLUSIONS	93
G. ACKNOWLEDGEMENTS.....	93
H. REFERENCES	93
I. FIGURES AND TABLES.....	97
3. RELATING THE OUTCOME OF HCV INFECTION WITH DIFFERENT HOST SNP POLYMORPHISMS IN A MAJORCAN COINFECTED HCV-HIV POPULATION TREATED WITH PEGIFN-RIB.	101
A. ABSTRACT	102
B. INTRODUCTION.....	103
C. MATERIALS AND METHODS.....	106
D. RESULTS	109
<i>Baseline characteristics of the subjects</i>	109
<i>Frequencies of the polymorphisms</i>	109
<i>Response to pegIFN-RBV treatment</i>	110
E. DISCUSSION	111
F. ACKNOWLEDGEMENTS.....	113
G. REFERENCES	113
H. FIGURES AND TABLES.....	118
5. DISCUSIÓ	123
1. ESTUDI DEL GENOMA VÍRIC <i>versus</i> DANY HEPÀTIC.....	125
2. ESTUDI DELS POLIMORFISMES EN PACIENTS <i>versus</i> RESPOSTA AL TRACTAMENT	131
6. CONCLUSIONS	135

7. BIBLIOGRAFIA	139
1. BIBLIOGRAFIA PERSONAL	151

1. EL VIRUS DE L'HEPATITIS C (VHC)

El virus de l'hepatitis C és la única espècie del gènere *Hepacivirus* dins la família *Flaviviridae*. A mitjans dels anys 70, es va informar que, després d'algunes transfusions sanguínies, hi havia pacients on apareixia una forma d'hepatitis que es va definir com a no A i no B (Feinstone *et al.*, 1975). Malgrat això, no va ser fins a finals dels anys 80 quan es varen aconseguir identificar les primeres seqüències del virus (Choo *et al.*, 1989). Aquest virus és la causa de problemes, tan aguts com crònics, en el fetge, tals com l'hepatitis crònica, la cirrosi o el carcinoma hepatocel·lular.

A. Estructura del VHC

El VHC presenta un genoma d'RNA de cadena simple i amb polaritat positiva (pot ser traduït directament per la cèl·lula hoste), d'aproximadament 9,6Kb, que codifica una poliproteïna d'uns 3000 aminoàcids.

El genoma del virus consisteix en un únic marc obert de lectura (ORF), el qual està franquejat per dues regions (5' i 3') no codificants i altament conservades evolutivament. La poliproteïna viral és processada per proteases de l'hoste i del virus, produint l'alliberació de 3 proteïnes estructurals i 7 no estructurals. Les proteïnes estructurals (la proteïna *core* i les dues proteïnes que configuren l'envolta) són tallades per peptidases de l'hoste, mentre que les proteïnes no estructurals (*p7*, *NS2*, *NS3*, *NS4A*, *NS4B*, *NS5A* i *NS5B*) es processen mitjançant proteases del propi virus (Lindenbach i Rice, 2005; Dubuisson, 2007) (Figura 1).

En la regió 5' no codificant hi ha un lloc d'unió al ribosoma, on s'uneix la subunitat ribosomal 40S i on s'inicia la traducció del genoma víric. Tal com s'ha esmentat abans, la poliproteïna és processada per proteasas de l'hoste i del virus, de manera cotraduccional i posttraduccional a nivell de la membrana del reticle endoplasmàtic per tal d'obtenir les 10 proteïnes madures.

La proteïna *core*, d'uns 191 aminoàcids, és una proteïna d'unió a RNA que forma la nucleocàpsida viral, i que intervé en nombroses funcions cel·lulars de l'hoste. Té una gran quantitat de residus hidrofòbics altament conservats, necessaris per un correcte plegament de la proteïna i per mantenir la seva estabilitat. Encara no s'ha elucidat exactament com s'assembla la nucleocàpsida, però s'ha pogut localitzar en diferents parts de la cèl·lula: en el reticle endoplasmàtic, el mitocondri, el citosol i en la superfície de gotes lipídiques, entre

d'altres (Penin *et al.*, 2004; Lindenbach i Rice, 2005; He i Tan, 2011). Una de les vies metabòliques amb les que interacciona la proteïna *core* és el metabolisme lipídic, afavorint, algunes variants, el desenvolupament d'esteatosi (acumulació de material lipídic en el fetge). L'esteatosi és un dels factors que contribueixen en gran mesura a l'avanç de la fibrosi hepàtica (Khan *et al.*, 2010).

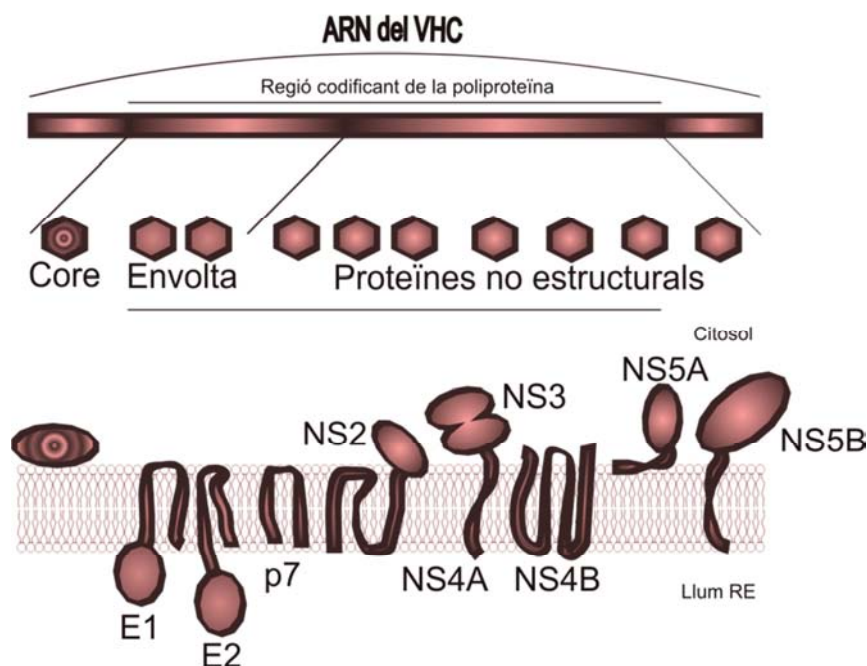


Figura 1. Estructura del genoma del VHC.

La proteïna *NS5B* està formada per 591 aminoàcids i és una RNA polimerasa dependent d'RNA, per tant, és l'element catalític de la maquinària replicativa del virus. Aquesta proteïna s'ha convertit en un objectiu clau de les noves teràpies contra el VHC, encara que la seva variabilitat genètica podria limitar la seva eficàcia degut a l'existència de variants resistents (Beaulieu, 2009)

Cicle vital del VHC

El virus de l'hepatitis C circula en diferents conformacions pel sèrum dels individus afectats (Penin *et al.*, 2004):

- i. Virions (o partícules víriques) units a lipoproteïnes de baixa densitat (LDL) o de molt baixa densitat (VLDL), que representen la fracció infecciosa.
- ii. Virions units a immunoglobulines.
- iii. Virions lliures.

El cicle vital del VHC comença amb les partícules víriques infeccioses circulant pel sèrum, acomplexades amb partícules lipoproteiques. Aquests complexos s'uneixen a les cèl·lules a través de receptors específics. Els complexos entren a dins les cèl·lules mitjançant una endocitosi amb receptors, modulada per vesícules recobertes de clatrina, i, amb la fusió del virió amb la vesícula endosomal, es produeix l'alliberació de l'RNA víric. L'RNA de polaritat positiva actua directament com a missatger sense la necessitat de cap proteïna accessòria per iniciar la infecció.

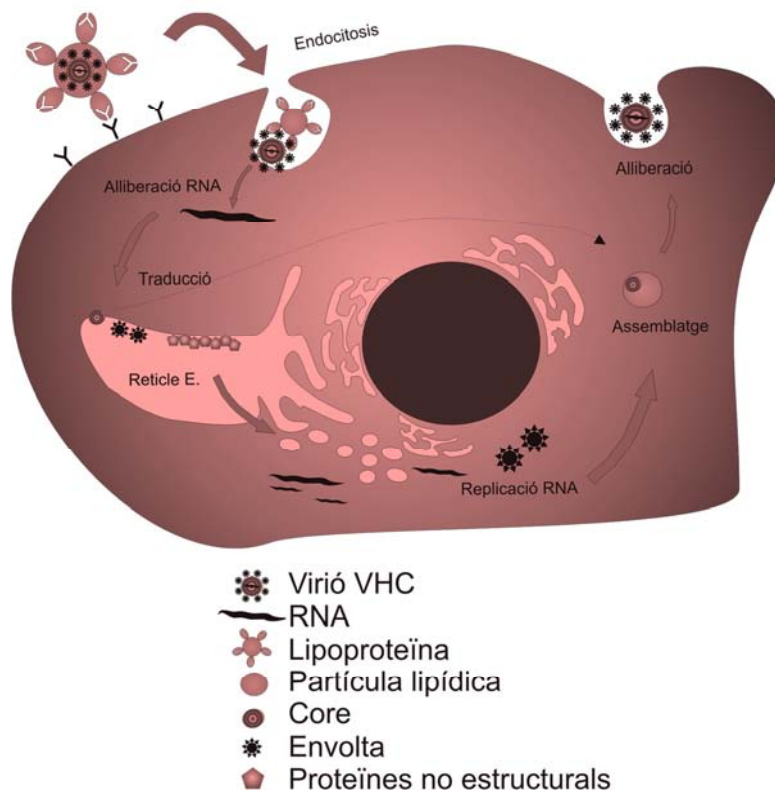


Figura 2. Esquema del cicle vital del VHC.

En aquest moment, les proteïnes traduïdes es mouen al compartiment cel·lular adequat a la seva funció. La proteïna *core* s'associa a gotes lipídiques, mentre que les glicoproteïnes de l'envolta i les proteïnes no estructurals es dirigeixen al reticle endoplasmàtic (RE). Les proteïnes no estructurals indueixen una expansió de la membrana del RE per tal de crear el lloc de la replicació de l'RNA, anomenat xarxa membranosa. La replicació produeix un RNA intermediari de polaritat negativa que serveix per crear nombroses noves cadenes d'RNA de polaritat positiva. Algunes d'aquestes noves cadenes entren al procés de traducció, algunes serveixen per produir l'RNA intermediari de polaritat negativa i d'altres participen en l'assemblatge de nous virions.

El procés d'assemblatge és el menys conegut, produeix noves partícules víriques aconseguint que molècules de la proteïna *core* associades a partícules lipídiques envoltin les molècules d'RNA finalitzant amb la unió a les glicoproteïnes de l'envolta (He i Tan, 2011).

S'ha de tenir en compte que aquest cicle vital s'ha explicat de manera molt breu i simplificada, i que cada fase és un esdeveniment que involucra a multituds de factors virals i de l'hoste, alguns coneguts, i d'altres no (Figura 2).

B. Genotips i variabilitat

El genoma del VHC presenta una gran variabilitat de seqüència degut, en gran mesura, a la falta de la prova de lectura de la seva RNAsa dependent d'RNA (*NS5B*).

La variabilitat genètica del VHC va ser descoberta a principi dels anys 90, moment en el qual, es varen haver de desenvolupar mètodes per classificar les diferents variants. Els virus de l'hepatitis C finalment s'agruparen en sis grups o *clusters* ben definits, anomenats genotips, els quals, a la vegada, es divideixen en subtipus (Figura 3). Els genotips es diferencien en un 31-33% a nivell de nucleòtids, i els subtipus en un 20-25%. Malgrat aquesta variabilitat, totes les soques comparteixen la mateixa longitud i els mateixos gens. Les regions més conservades són els extrems 5' i 3' seguides de la regió *core*. Pel contrari, les regions més variables corresponen a les proteïnes de l'envolta (Simmonds *et al.*, 2005).

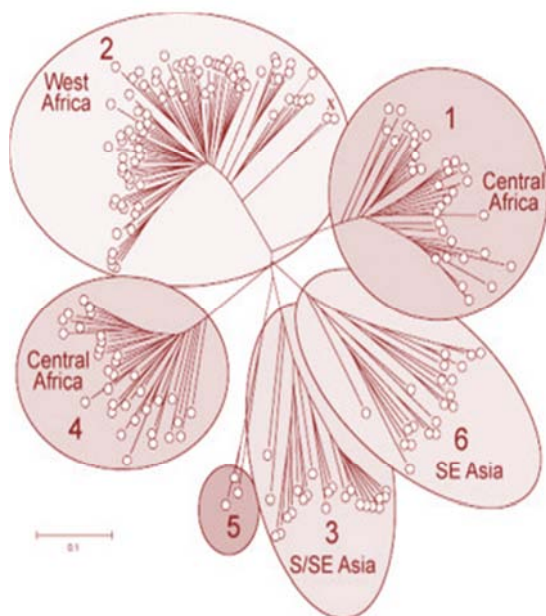


Figura 3. Arbre evolutiu de les seqüències de la regió *NS5B* (segons Simmonds *et al.*, 2005).

Aquesta variabilitat i la ràpida evolució del virus té una sèrie de conseqüències pràctiques. Per exemple, pot permetre conèixer el començament i evolució d'un brot infeccios a partir de l'anàlisi filogenètica dels virus que estan infectant els diferents pacients. És a dir, es pot conèixer si tots els pacients es varen contagiar en un mateix brot i, si algun d'ells, és el focus d'infecció inicial (González-Candelas *et al.*, 2003).

Altres anàlisis relacionades amb les diferències existents entre els diferents genotips, han determinat que poden presentar diferents pronòstics o tendències en la malaltia. Per exemple, s'ha assenyalat que la proteïna *core* codificada per virus del genotip 3 indueix esteatosis més eficientment que la produïda pels virus de genotip 1 (Khan *et al.*, 2010). Per un altre banda, Poynard *et al.* (2003) també trobà que el genotip 3 està possiblement associat amb un major desenvolupament de fibrosi. Aquests treballs, entre d'altres, indiquen que estudiar la variabilitat dins del virus de l'hepatitis C és important per fer un pronòstic del curs de la malaltia.

C. Prevalença i distribució geogràfica

L'Organització Mundial de la Salut (OMS) estima que fins a un 3% de la població mundial ha estat infectada pel virus de l'hepatitis C, això significa que existirien uns 170 milions de portadors del virus (Thomson i Finch, 2005).

La prevalença del virus de l'hepatitis C va des de menys de l'1% al Nord d'Europa fins a més del 2% de mitjana en el Nord d'Àfrica. La prevalença més elevada detectada fins al moment la podem trobar a Egipte (al voltant d'un 15-20%) (Alter, 2007). El desenvolupament de mètodes més eficients de cribratge dels donants de sang, ha permès que els casos d'infecció del VHC per transfusions sanguínies siguin extremadament rars. La prevalença del VHC s'ha incrementat en les darreres dècades, com a conseqüència de la propagació del virus a través de xarxes d'administració de drogues via parenteral (ADVP) (Pybus *et al.*, 2005; Thomson i Finch, 2005). En la Figura 4, es pot veure un mapa de la prevalença estimada de la infecció a Europa.

Les freqüències relatives dels diferents genotips del VHC en les diferents poblacions no són uniformes (Figura 5), reflectint canvis segons els grups poblacionals i l'edat i via de transmissió. Per tal cosa, la distribució dels genotips i subtipus ha anat canviant al llarg del temps. En els països industrialitzats, el subtip **1a** es troba, generalment, en individus joves compartint com a factor de risc ADVP. Per una altra banda, el subtip **1b** és més comú en

persones més majors, amb un historial de transfusions sanguínies (Pybus *et al.*, 2005). A Espanya, la prevalença del virus és similar a altres països industrialitzats, sent el genotip **1** el més freqüent seguit del genotip **3** (Bruguera i Forns, 2006; Esteban *et al.*, 2008), encara que la distribució dels subtipus presenta algunes diferències amb altres països europeus (López-Labrador *et al.*, 1997), amb una elevada influència d'infeccions definides pels subtipus **1b**. La distribució en la nostra comunitat és similar a la resta del país (Cifuentes *et al.*, 2004).

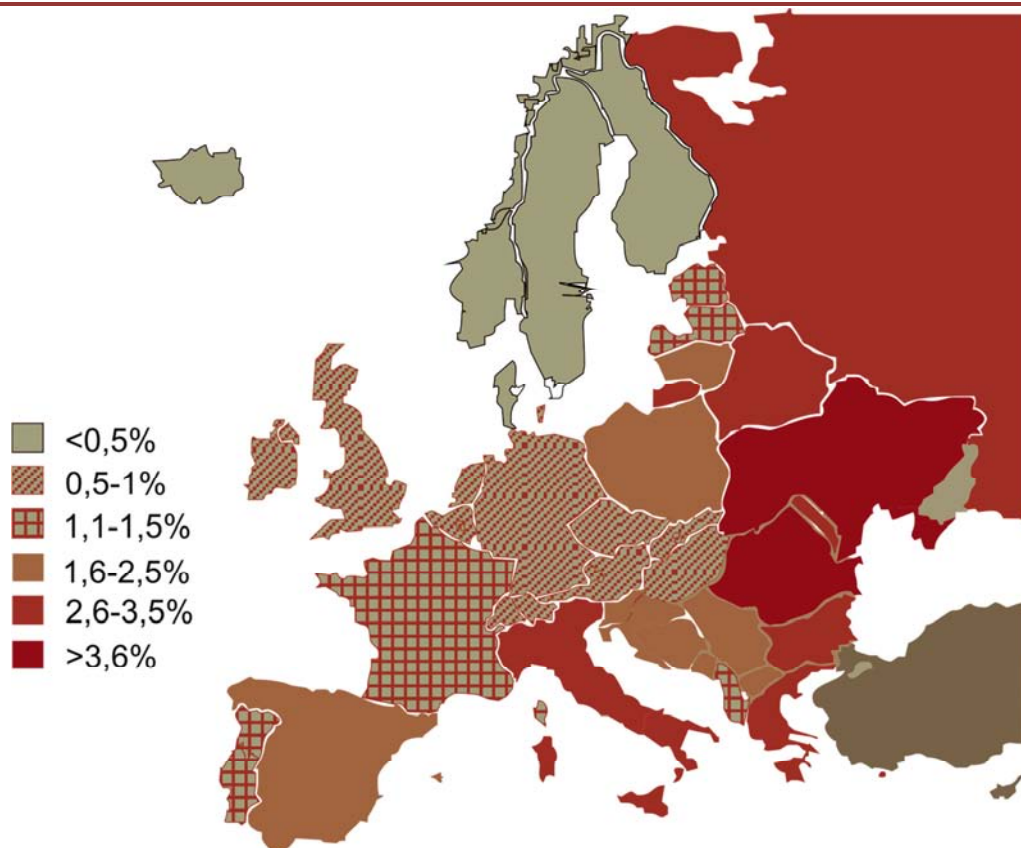


Figura 4. Mapa de la prevalença de la infecció pel VHC a Europa (adaptat d'Esteban *et al.*, 2008).

Els estudis més recents de cohorts a Espanya mostren un descens relatiu de les freqüències dels genotips **1b** i **2**, a la vegada que augmenten les infeccions causades per virus del subtip **1a**, probablement degudes a un increment d'infeccions d'aquest subtip entre els usuaris de drogues d'ús via parenteral (Jiménez-Hernández *et al.*, 2007). Els darrers 30 anys sembla que han ocorregut a Espanya dues grans epidèmies del VHC separades. Primer, una involucrada amb l'expansió dels genotips **1a** i **3** i, més recentment, una relacionada amb l'expansió del genotip **4** (Echevarría *et al.*, 2006).

Com a dada positiva cal dir que la incidència de la infecció pel virus de l'hepatitis C ha disminuït més d'un 50% durant la darrera dècada, la qual cosa reflecteix una reducció significativa de les exposicions de risc (Gane, 2011).

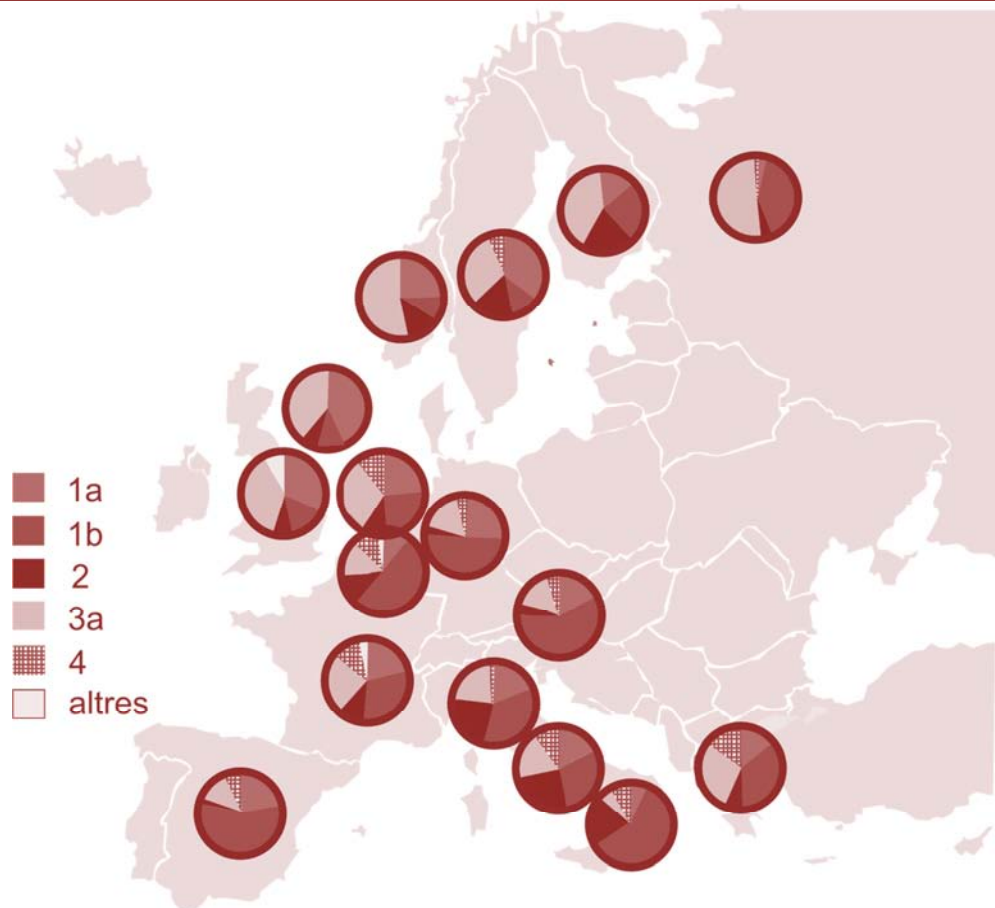


Figura 5. Distribució dels diferents genotips del virus de l'hepatitis C a Europa (adaptat d'Esteban *et al.*, 2008)

2. PATOLOGIA DEL VHC

A. Història natural de la infecció

La transmissió del virus de l'hepatitis C es produeix bàsicament a través de la sang o derivats sanguinis, sent gran part de la seva dispersió deguda a transfusions sanguínies i al fet de compartir agulles i xeringues contaminades. Avui en dia, la transmissió del virus a través de l'ús de drogues d'administració per via parenteral és la més freqüent. Malgrat això, alguns pacients adquireixen la malaltia sense cap exposició coneguda, ni a transfusions ni a drogues intravenoses (Poynard *et al.*, 2003).

La història natural de la infecció pel VHC encara esdevé complicada d'explicar en la seva totalitat. En gran mesura és complicat establir-la perquè l'inici de la infecció és difícil de determinar i, perquè, en la majoria de casos la infecció progressa cap a un estat crònic, la durada del qual és molt variable. El marge de temps que transcorre entre la infecció crònica i una resolució final definida és molt ampli i pot anar de 15 fins a uns 40 anys. Una altra característica és que, normalment, la infecció crònica, incloent moltes vegades l'inici del procés de cirrosi, transcorre sense símptomes (Buti i Esteban, 2003).

Una vegada s'ha fet crònica la malaltia, la successió d'esdeveniments sol començar per un avanç progressiu de la fibrosi, seguida d'un procés de cirrosi, primer compensada i després descompensada, acabant finalment en un hepatocarcinoma.



Figura 6. Representació gràfica de l'evolució natural de la infecció pel VHC (adaptada d'Asselah *et al.*, 2009).

És necessari tenir en compte uns quants ítems per tal d'explicar d'alguna manera el curs de la malaltia. El primer és establir quin és la ràtio d'eliminació (o aclariment) espontània del virus després de la infecció. Aquest percentatge està al voltant del 15-30% del total d'infectats adults, contrastant amb un grau de cronificació d'aproximadament 75-85%, segons diferents estudis (Figura 6). Una ràtio més elevada d'aclariment la trobem a infants i dones joves, arribant fins al 45% del total d'infectats en aquest rang d'edats. A dos estudis realitzats en dones embarassades que s'infectaren amb preparacions d'immunoglobulines anti-D, es detectà que la meitat eliminaren espontàniament el virus, mentre que l'altra meitat va fer crònica la malaltia (Seeff, 2009; Rauch *et al.*, 2010a). Aquest grup d'edat també presenta les ràtios més baixes de desenvolupament de cirrosi i hepatocarcinoma. Una altra informació important que s'hauria de conèixer, és quin resultat de la malaltia es pot anticipar entre aquelles persones que no han eliminat espontàniament el virus. Relacionada amb aquesta informació, trobaríem quines variables es poden

contemplar com a factors de risc que ajudessin a predir la progressió de l'hepatitis C crònica. Finalment, s'hauria de conèixer quina és la patogènesi de la malaltia persistent i els mecanismes pels quals progressa en un sentit o un altre (Seeff, 2009).

Poynard *et al.* (2003) revisaren tots els factors coneguts que poguessin estar associats amb el desenvolupament de la fibrosi i, per tant, de la malaltia en general. Ateses les dades que mostren una eliminació espontània del virus en infants i dones joves, pareix ser que l'edat en el moment de la infecció és un factor clau en la probabilitat d'un aclariment espontani del virus. Malgrat això, l'edat també podria ser un indicador indirecte de la vigor de la resposta immune (Alter, 2006a). El sexe masculí té un pronòstic pitjor que el femení en aquesta infecció, així com el fet de ser consumidor habitual d'alcohol.

Un altre grup de variables que poden intervenir en el progrés de la malaltia són les relacionades directament amb el virus, com per exemple el seu genotip o la càrrega viral que presenten, emperò les evidències encara són escasses. Per una altra banda, també s'ha observat que pacients amb els nivells de transaminases (GOT i GPT) baixos presenten una fibrosi significativament més baixa que els pacients amb aquests enzims elevats (Seeff, 2009). Un altre enzim hepàtic, la gamma glutamiltranspeptidasa (GGT) s'ha trobat en quantitats més baixes en pacients que responen més bé al tractament, i més elevades en pacients amb esteatosis hepàtica i fibrosi avançada (Kau *et al.*, 2008). La resistència a la insulina i l'obesitat també suposen un major risc de fibrosi (Thomas *et al.*, 2009).

Estudis recents posen de manifest que el metabolisme lipídic es veu molt alterat en el transcurs de la infecció pel VHC. Per exemple, les dades aportades per Piver *et al.* (2010) mostren que els nivells de triglicèrids i colesterol en pacients amb hepatitis C crònica, són més baixos que en pacients que han eliminat espontàniament la malaltia. En el mateix sentit, també indiquen que pacients amb els nivells de triglicèrids més alts, mostren una millor resposta al tractament.

El control de la infecció pel virus de l'hepatitis C és mediat immunològicament, fet que es fa evident en l'aparició de cèl·lules T específiques pel virus. La persistència de la infecció ocorre quan la resposta innata es troba disminuïda, de manera que els efectes antivirals es troben reduïts. La inducció de cèl·lules T produïda per l'interferó- γ (IFN- γ) ajuda a eliminar la infecció, mentre que una resposta pobre d'aquest interferó permet que continuï la replicació viral i els seus efectes amb l'alliberació de citosines (Seeff, 2009).

D'aquesta manera, l'hepatitis C crònica no és conseqüència de la destrucció directa de les cèl·lules hepàtiques per part del virus, sinó que és resultat d'una resposta immune intermèdia suficientment gran per induir la destrucció de les cèl·lules hepàtiques i la formació de fibrosi, però no suficientment gran per eliminar els virus (Poynard *et al.*, 2003).

Els estudis de la relació entre la persistència viral i la disfunció en les cèl·lules T CD4 i CD8 indiquen que els pacients que s'han recuperat de l'hepatitis tenen una resposta d'aquests tipus de cèl·lules molt més vigorosa i reactiva que els pacients que han cronificat la malaltia, és a dir, es tracta d'una major funcionalitat, no d'un canvi en el nombre de cèl·lules. En una primera resposta l'actuació de les cèl·lules T és semblant en els dos grups, emperò en els pacients que han fet crònica la malaltia, les cèl·lules T funcionals pareix que progressivament queden exhaustes. Per tal cosa, aquesta observació seria una conseqüència i no la seva causa del pas a crònic. Es troben altres alteracions en la resposta immune que difereixen entre els pacients que eliminen el virus i els que no ho fan. Hi ha una sèrie de cèl·lules T reguladores que inhibeixen l'acció de les CD8, mentre que una depleció d'aquestes cèl·lules reguladores produeix una millor actuació de les CD8. Terapèuticament es podria intentar eliminar aquestes cèl·lules reguladores, però es correria el risc de causar un dany hepàtic sever induït per les cèl·lules CD8 o produir un fenomen autoimmune. La disfunció en les cèl·lules CD8 afecta particularment a la seva habilitat per produir interferó α (IFN- α), la qual cosa és un esdeveniment clau en la resposta antiviral innata. La supressió de la producció de l'IFN- α està induïda per diferents proteïnes virals, per tant el virus emprava proteïnes pròpies per escapar de la resposta immune (Alter, 2006a).

Finalment, cal a dir que la situació esdevé més complicada respecte al desenvolupament de la malaltia. Entre els experts hi ha dues corrents de pensament contràries, una indica que la progressió cap a un estadi terminal de la malaltia hepàtica o cap a un hepatocarcinoma, és un fet inevitable, sempre que passi el suficient temps. Per altra banda, trobem el punt de vista dels que pensen que són una minoria els pacients infectats que arriben a una situació de malaltia hepàtica terminal (Buti i Esteban, 2003).

B. Diagnòstic

Un dels fets que més dificulta la possibilitat d'establir la història natural de la infecció és que el diagnòstic de la malaltia, moltes vegades es realitza a través d'anàlisis bioquímiques fetes a l'atzar o per causa d'altres malalties. En aquestes anàlisis es poden

trobar alteracions dels enzims hepàtics, que poden induir a pensar que estam davant una hepatitis que es podria confirmar o no per estudis serològics (Buti i Esteban, 2003).

Les proves diagnòstiques es poden agrupar en assaigs serològics mitjançant anticossos i proves moleculars per detectar partícules víriques. Els anticossos contra el VHC són detectats emprant immunoassaigs enzimàtics que són molt específics i sensibles. Els darrers immunoassaigs desenvolupats poden detectar anticossos dins les primeres 4-10 setmanes de la infecció. En contra ens trobam que els pacients immunosuprimits, per exemple els diagnosticats amb el VIH, poden presentar falsos negatius. Les proves qualitatives per detectar l'RNA víric estan basades en la tècnica de la PCR, i tenen un límit de detecció de menys de 100 còpies d'RNA del VHC per mil·lilitre de sèrum. És una tècnica molt sensible i utilitzada quan l'ús d'anticossos suposa algun problema o dona resultats contradictoris (Poynard *et al.*, 2003).

C. Tractament

Ara per ara, no existeix cap vacuna o tractament definitiu contra el VHC, tot i que, sabent que existeix un procés d'eliminació de manera natural del virus, és d'esperar que algun dia s'esbrinin els mecanismes que el duen a terme. De manera teòrica, la infecció pel virus hauria d'activar les vies d'activació dels interferons de classe I, emperò l'RNA d'aquest virus codifica per proteïnes específiques que inhibeixen la inducció dels interferons. Aquest fet permet al virus evadir les accions antivirals de l'hoste (Asselah *et al.*, 2009).

Durant molts d'anys el tractament més eficaç emprà l'interferó, produint una resposta viral sostinguda del 20%. Es defineix com a resposta viral sostinguda (RVS) que l'RNA víric sigui indetectable en la sang perifèrica del pacient després de 24 setmanes d'haver acabat el tractament. Durant el tractament, la càrrega viral del VHC és monitoritzada. Els pacients que no responen des del començament no modifiquen la càrrega viral, mentre els que la disminueixen significativament de forma primerenca (reducció de més de 2 log) o bé negativitzen ràpidament la càrrega viral (a les 4 setmanes), tenen un millor pronòstic d'aconseguir una RVS. L'ús conjunt d'interferó amb ribavirina i el posterior ús d'interferó polietilenglicolat (IFN-peg) varen suposar un gran avanç en el tractament de la malaltia produïda pel VHC (Thomson i Finch, 2005; Asselah *et al.*, 2010). Actualment, aquest tractament es considerat el tractament estàndard (o SOC, *Standard of Care*), encara que es considera que falla en un 50% dels casos, de mitjana, a més de presentar efectes secundaris complicats, i ser un tractament d'alt cost econòmic. Per tant,

seria interessant poder conèixer *a priori* quins són els mecanismes que s'amaguen darrera els pacients que no responen al tractament, per conèixer quines possibilitats es té en cada cas de respondre i de quina manera s'ha d'administrar aquest tractament (Suppiah *et al.*, 2011). Es poden diferenciar dos tipus d'elements associats a pacients que no responen al tractament. El primer tipus inclou els factors virals, dels quals, el genotip del virus potser sigui el més rellevant. En pacients infectats per VHC del genotip 2 o 3, la RVS arriba al 80%, mentre només un 50% dels pacients infectats per virus del genotip 1 assoliran la RVS. Els pacients infectats amb el genotip 4 es troben en un terme intermedi, un 65% de mitjana mantindran una RVS. La variabilitat dels virus també suposa un element a tenir en compte. La possibilitat de formar quasiespècies, podria permetre al virus generar variants amb diferent sensibilitat al tractament, per exemple estabilitzant el complex de proteïnes víriques. Dins aquest grup de factors, també es podria incloure la càrrega viral; diversos estudis han relacionat càrregues virals elevades amb una menor sensibilitat al tractament. De l'aplicació d'aquests coneixements ha sorgit el fet que a un pacient infectat amb un VHC del genotip 1, baixa càrrega viral i resposta ràpida a les 4 setmanes se l'hi administra el tractament durant 24 setmanes, mentre que el tractament a un pacient infectat amb el genotip 3, elevada càrrega viral i sense resposta ràpida, hauria d'administrar-se durant 48 setmanes. Malgrat això, pacients infectats amb el mateix genotip i amb càrregues virals semblants, poden mostrar respostes al tractament diferents, pel que s'entreveu la importància d'un segon tipus de factors (Rauch *et al.*, 2010a). Aquest inclou característiques pròpies de l'hoste, o independents del virus, com per exemple l'edat avançada, pertànyer al gènere masculí, la presència de cirrosi o esteatosi hepàtica, la resistència a la insulina, la presència de diabetis o el pes, tots ells relacionats amb una menor RVS. La coinfecció amb el virus de la immunodeficiència humana (HIV) o amb el virus de la hepatitis B (VHB) està també associada a una menor RVS (Poynard *et al.*, 2003; Heathcote, 2007; Asselah *et al.*, 2010).

La identificació de nous marcadors moleculars poden ajudar a discriminar pacients que no respondran al tractament, tal com es veurà al punt 5 d'aquesta introducció. De manera semblant, estudiar els diferents patrons d'expressió gènica que poden existir entre pacients que responen i els que no, pot ajudar a esbrinar la diferència en la eficàcia del tractament. Aquests patrons diferents d'expressió es poden trobar tant en el fetge com a cèl·lules sanguínies, on seria més interessant el seu estudi perquè no es necessitarien proves invasives (Asselah *et al.*, 2010).

Degut a les limitacions del tractament actual, sobretot per la baixa resposta d'alguns genotips i els efectes secundaris que té, es continuen investigant nous tractaments, o noves pautes dels tractaments existents, per exemple emprant ribavirina a dosis molt altes administrades de forma escalada. Un altre camp d'estudi ha estat l'ús de molècules antisentit dirigides a inhibir el lloc d'unió a ribosoma de la regió 5' del virus (Alter, 2006a).

Gane (2011) presenta un compendi de les noves teràpies que s'estan posant a punt per tractar aquesta malaltia. Hi ha cinc passes en la replicació del virus que poden ser objectius d'una teràpia antiviral directa. La primera inclou la unió del virus a l'hepatòcit a través de receptors de superfície. La traducció i processament de la poliproteïna són també nous objectius, seguits de l'assemblatge i maduració del virus i la seva posterior alliberació de la cèl·lula. Fins ara, els estudis més recixits van dirigits a inhibir les proteases del virus (*NS3*) i el complex de l'VHC polimerasa (*NS5A*). En els darrers 5 anys més de 90 inhibidors de la proteasa i la polimerasa vírica han entrat en la fase de desenvolupament clínic, encara que molts d'ells s'han hagut d'abandonar donada la seva toxicitat o manca d'eficàcia. Dos inhibidors de proteases (*Boceprevir* i *Telaprevir*) estan entrant en les darreres fases de desenvolupament, i s'estan obtenint les primeres dades de la que s'ha anomenat "teràpia triple" (IFN-peg, ribavirina i els inhibidors de proteases) que han estat encoratjadores. En la majoria de casos s'ha augmentat la RVS de pacients infectats amb virus del genotip 1 del 50 al 70%, alhora que s'ha pogut escurçar la durada del tractament (Asselah *et al.*, 2010). Una altra via d'estudi, emprant com a base el tractament actual pel VIH, és l'administració simultània, via oral, de diferents composts que actuïn en diferents fases de la replicació del virus.

3. COINFECCIÓ VHC-VIH

A. VIH com a modificador de la infecció pel VHC

Les dades de que disposem suggereixen que els pacients coinfectats amb el virus de l'hepatitis C i el virus de la immunodeficiència humana presenten un major risc de problemes hepàtics i una més ràpida evolució de la malaltia que els pacients monoinfectats amb VHC (Matthews i Dore, 2008). Els pacients coinfectats tenen una càrrega viral més elevada, una inflamació més greu i una progressió més accelerada cap a cirrosi que els pacients només infectats pel VHC. Per tant, la mortalitat relacionada amb problemes hepàtics és més gran que en pacients monoinfectats (Alter, 2006a).

En un estudi prospectiu fet a Espanya, en una cohort de pacients coinfectats, s'estima que després d'un període de 3 anys, el 44% dels pacients, progressaran en un o més estats de la fibrosi en un període curt de temps (Macías *et al.*, 2009). A més a més, la progressió de la malaltia per infecció del virus de l'hepatitis C, és més pronunciada en pacients amb recomptes baixos de CD4+ (Reiberger *et al.*, 2010); encara que no es va poder establir si les variacions en el desenvolupament de la malaltia hepàtica en pacients immunosuprimits, estan relacionades amb diferències genètiques del virus de l'hepatitis C (Gigou *et al.*, 2001). En quant a la variabilitat del virus, en una cohort de pacients infectats amb VHC del subtip **1b**, hi ha una correlació entre la similitud genètica de les regions *core* i *NS5B* del virus i les similituds en la progressió de la fibrosi en pacients no immunosuprimits, però no en pacients amb el fetge trasplantat i immunosuprimits (López-Labrador *et al.*, 2004; 2006). Al mateix temps, s'ha observat una reducció de la diversitat genètica del VHC en pacients coinfectats amb VIH, si es compara amb pacients monoïnfectats, suggerint una reducció de la pressió immunoselectiva (Shuhart *et al.*, 2006; López-Labrador *et al.*, 2007).

És important destacar el fet que els pacients coinfectats també han de rebre tractament pel VIH. Aquests tractaments antiretrovirals estan sovint associats a un augment del nivell de transaminases, cosa que indica una modificació de l'activitat hepàtica, i que poden estar relacionats amb un augment de la fibrosi (Poynard *et al.*, 2003). Alguns estudis indiquen que hi ha de 2 a 10 vegades més de probabilitats que un pacient sofreixi hepatotoxicitat quan és tractat amb els antiretrovirals d'alta activitat (TARGA) (Thomson i Finch, 2005).

B. Prevalença i distribució de la coinfecció VHC-VIH

Mentre la prevalença de coinfecció VHC i el virus de la immunodeficiència humana-1 (VIH-1, família *Retroviridae*, gènere *Lentivirus*) en la població general es troba al voltant del 2%, en pacients amb una història d'ADVP, aquesta prevalença puja fins a un 70-95% (Alter, 2006b). La coinfecció VHC-HIV, és un problema clínic de nivell mundial, i està relacionada amb una mortalitat més elevada que si la comparem amb la que produeix cada virus individualment (Rotman i Liang, 2009).

4. DANY HEPÀTIC

A. Fibrosi

Encara que el virus de l'Hepatitis C no es considera citopàtic per sí mateix, l'efecte patogènic de la infecció crònica és la generació de fibrosi hepàtica amb el temps (McCaughan i George, 2004). La fibrosis es converteix progressivament en cirrosi del fetge i, conseqüentment, condueix a un estat terminal de la malaltia.







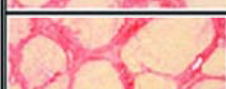
La definició més àmpliament acceptada de la fibrosi és l'acumulació en excés de col·lagen i altres materials de la matriu extracel·lular, degut a la formació de noves fibres. La fibrogènesi és un procés dinàmic complex relacionat amb la necroinflamació i l'activació de cèl·lules estrellades hepàtiques. Per avaluar els canvis que es produeixen en el transcurs de la malaltia o en l'eficàcia del tractament, és necessari un coneixement precís de l'estat de la fibrosi. Per avaluar el contingut de col·lagen de manera directa és necessari realitzar una biòpsia del teixit, a partir del qual es podria saber quina quantitat de col·lagen hi ha mitjançant mètodes bioquímics. Aquests mètodes impliquen la destrucció del teixit i la impossibilitat de fer anàlisis posteriors, per tant la majoria de tècniques s'han basat en l'avaluació histològica, emprant la tinció tricròmica o de reticulina i més recentment, la tinció *Sirius Red* (Asselah *et al.*, 2009; Germani *et al.*, 2010).

S'han desenvolupat diferents sistemes per tal de classificar, de forma objectiva i sistemàtica els estadis de la fibrosi hepàtica. Alguns dels més emprats són el Knodell, l'Ishak i l'escala Metavir, encara que, tots tenen pros i contres (Theise, 2007). En la Taula 1, es poden veure els detalls histològics d'un teixit hepàtic i la seva equivalència en les diferents escales de mesura de la fibrosi.

L'escala Metavir es va crear per avaluar la fibrosi específicament en pacients infectats pel virus de l'hepatitis C. En la Figura 7, es pot veure de forma esquematitzada la relació entre la morfologia del teixit hepàtic i els estadis de l'escala Metavir.

Avui en dia les biòpsies hepàtiques segueixen constituint *l'estàndard d'or* en l'avaluació de la fibrosi hepàtica, encara que algunes veus discordants indiquen que hi pot haver interpretacions dolentes degut a errors de mostreig (Bedossa *et al.*, 2003). Per exemple, un estudi demostrà que dues mostres de teixit hepàtic histològicament normal obtingudes utilitzant diferent procediments, presentaven diferents patrons d'expressió gènica, malgrat ambdues presentessin una histologia normal (Asselah *et al.*, 2009).

Taula 1. Relació entre les diferents escales de classificació del grau de fibrosi i la histologia del teixit hepàtic (tinció *Sirius Red*). En la darrera columna, es pot observar el percentatge de fibrosi per àrea que s'observa en cada teixit, en funció del col·lagen trobat (adaptat de Standish *et al.* (2006), amb modificacions).

Histologia	Descripció	Escala Ishak	Escala Knodell	Escala Metavir	Percentatge Fibrosi
	Sense fibrosi (normal)	F0	F0	F0	1,9%
	Fibrosi en algunes àrees portals amb algun septe petit	F1	F0-F1	F0-F1	3,0%
	Fibrosi en la majoria d'àrees portals amb septes petits	F2	F1	F1	3,6%
	Fibrosi en la majoria d'àrees portals amb alguns ponts fibròtics	F3	F1-F3	F2	6,5%
	Fibrosi en la majoria d'àrees portals amb molts de ponts fibròtics	F4	F3	F3	13,7%
	Cirrosi incompleta	F5	F3-F4	F3-F4	24,3%
	Cirrosi	F6	F4	F4	27,8%

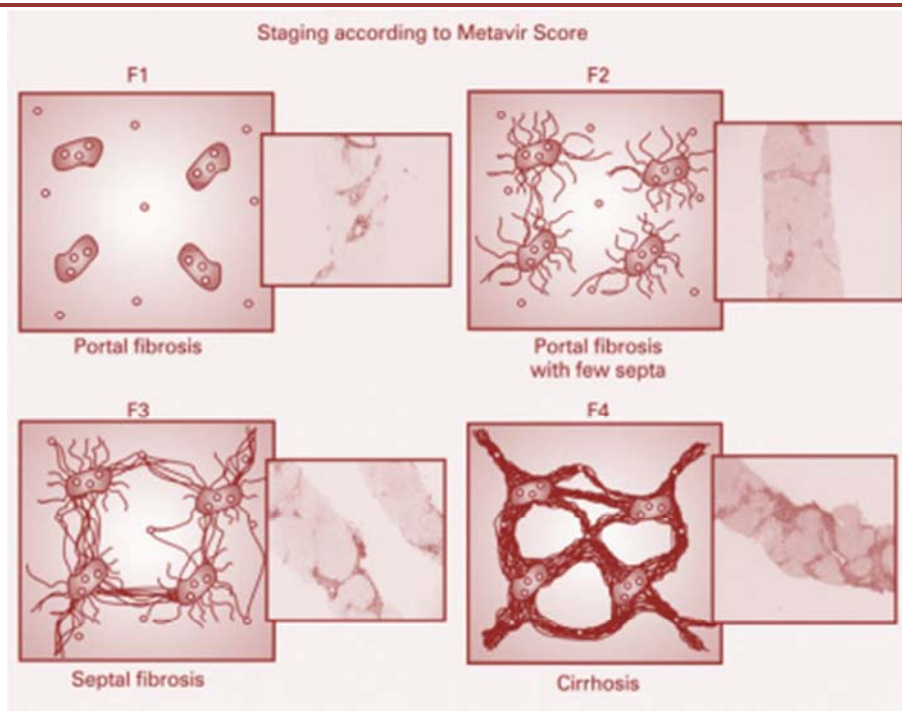


Figura 7. Esquema de la relació entre l'estructura i morfologia del teixit hepàtic i els diferents estadis de fibrosi avaluats amb l'escala Metavir (segons Asselah *et al.*, 2009).

B. Mesures no invasives del dany hepàtic

El desenvolupament de proves no invasives com a marcadors de la fibrosi hepàtica permet estimar el dany hepàtic sense la necessitat de dur a terme una biòpsia (Gressner *et al.*, 2007).

Entre els marcadors no invasius, trobem les tècniques d'imatge, els marcadors genètics i els marcadors sèrics. Una de les tècniques d'imatge més emprades és el FibroScan, ja sigui per sí mateix o en combinació amb altres tècniques. És una tècnica també anomenada “elastografia de transició”, que permet determinar la rigidesa del teixit hepàtic emprant un aparell similar a un ecògraf. Altrament, els marcadors genètics inclouen polimorfismes que poden tenir a veure amb el desenvolupament de la fibrosi. Finalment, hi ha nombrosos marcadors sèrics que poden servir per avaluar el grau de fibrosi, entre ells trobem l'índex de Forns, l'APRI i el FIB-4 (Ahmad *et al.*, 2011).

5. VARIABILITAT GENÈTICA DELS HOSTES

Com s'ha esmentat abans, el desenvolupament de la infecció amb el virus de l'hepatitis C pot ser de dos tipus. El primer s'assoleix en un 30% dels casos, quan la resposta immunològica innata aconsegueix un control permanent i definitiu de la infecció conegut com aclariment espontani. El segon, per tant, implica una cronificació de la malaltia quan aquests mecanismes de resposta fallen. Fins ara, les característiques víriques, demogràfiques i clíniques no han pogut explicar el perquè dels diferents patrons de desenvolupament de la malaltia (Rauch *et al.*, 2010b).

La via d'estudi que s'ha establert recentment es basa en la variabilitat genètica dels hostes. És a dir, en establir quines característiques genètiques del hostes poden donar un millor o pitjor pronòstic, tant a nivell de dany hepàtic com de resposta al tractament. En tot cas, seria interessant avançar el pronòstic de la malaltia per saber *a priori* si és necessari administrar un tractament més o menys agressiu.

Els primers estudis genètics involucraren polimorfismes en gens que estaven relacionats amb el control de la infecció del VHC, com per exemple, antígens de leucòcits, receptors de certes immunoglobulines, gens estimulats per interleucines o interferons, etc. Algun estudi va donar resultats positius, però aquests no foren reproduïbles. Recentment, amb la posada a punt de les tècniques d'anàlisi del genoma complet (GWAS, *genome-wide*

association studies) han permès detectar regions que pareix ser són importants per determinar aquestes diferències. Els estudis basats en el GWAS empen mapes genètics, que cobreixen tot el genoma humà, per tal de trobar diferències al·lèliques entre casos i controls. Una diferència significativa indicaria que una determinada regió estaria involucrada en el desenvolupament d'un tret. La fortalesa d'aquesta tècnica seria l'habilitat no només de revelar gens que, en principi podrien estar relacionats amb el tret en qüestió, sinó també, d'aquells que no tendrien cap relació, de manera que s'afegiria un nou punt de vista en la patofisiologia o farmacologia d'una malaltia (Asselah *et al.*, 2009; Rauch *et al.*, 2010b).

A. IL28b (rs12979860)

A l'any 2009, Ge *et al.*, Suppiah *et al.* i Tanaka *et al.* publicaren, gairebé simultàniament, per primera vegada la contribució de polimorfismes d'un sol nucleòtid o SNPs, en una nova regió propera al gen de la interleucina 28b (IL28b), també coneguda com interferó lambda 3 (IFN- λ -3), en el desenvolupament de la malaltia causada pel VHC. Aquests tipus d'interferons interaccionen amb els seus receptors específics induint respostes antivirals. Alguns polimorfismes propers a aquest gen, podrien modificar l'expressió de IL28b i, per tant, la resposta al tractament (Rallón *et al.*, 2010). Els tres estudis esmentats abans, utilitzaren el mètode GWAS en persones infectades amb el genotip 1 i analitzaren la resposta al tractament amb interferó pegilat i conjuntament, en alguns casos, amb ribavirina. Rauch *et al.* (2010a) ampliaren l'estudi a pacients infectats amb VHC d'altres genotips (1, 2, 3 i 4).

El polimorfisme que trobà més fortament lligat a una diferent resposta fou descrit com a rs12979860 (Ge *et al.*, 2009). Es troba en el cromosoma 19 a 3kb cadena amunt del gen IL28b. L'estudi de Ge *et al.* assenyalà que el genotip CC presenta una taxa de resposta viral sostinguda dues vegades superior al genotip TT. Aquest SNP, degut a la seva desigual distribució (l'al·lel C és dues vegades més freqüent en individus amb ancestres europeus) va suposar una primera explicació de perquè els pacients amb ancestres africans responien pitjor al tractament que els pacients amb ancestres europeus. També es destacà que l'al·lel C, a més d'estar relacionat amb una millor resposta, contradictòriament, estava associat amb una major càrrega viral. Aquest estudi també va suggerir que l'al·lel C podria afavorir l'aclariment espontani del virus, ja que aquest al·lel es trobà en menor freqüència en els pacients que havien cronificat la infecció que en individus controls, fet confirmat a l'estudi de Thomas *et al.* (2009).

Per una altra banda, aquesta millora en la resposta al tractament que suposa el genotip CC, resulta més evident en els pacients infectats per virus dels genotips 1 i 4, els quals independentment del genotip que presenti el gen IL28b, sempre havien mostrat una resposta pitjor al tractament que els infectats per virus de genotips 2 i 3 (Rauch *et al.*, 2010a)

A la Figura 8 es pot observar una representació de les freqüències dels al·lels C i T del polimorfisme rs12979860 en població infectada amb el VHC en el món, s'ha de destacar que l'al·lel C està gairebé fixat a la zona d'Àsia, mentre l'al·lel T és el més freqüent en moltes zones d'Àfrica. A Europa trobam freqüències intermèdies.

Estudis realitzats amb pacients coinfectats amb el VIH, també donaren resultats semblants, es a dir, una millor resposta al genotip CC, en front els altres genotips (Rallón *et al.*, 2010; Rauch *et al.*, 2010b).



Figura 8. Freqüències gèniques del polimorfisme del gen IL28b (en clar s'assenyala l'al·lel C i en fosc, l'al·lel T) (segons Thomas *et al.*, 2009).

Li *et al.* (2010) investigaren si la variabilitat en aquest polimorfisme estava associada als menors nivells de LDLs que s'havien trobat en els pacients que havien cronificat la malaltia. Els resultats mostraren que els pacients amb el genotip CC presentaven un nivell significativament major d'aquestes lipoproteïnes. L'explicació podria ser que el genotip CC,

que és el que representa una millor resposta al tractament, podria estar associat a uns menors nivells d'interferó o una menor expressió dels gens activats per interferons. Això resultaria en un increment dels nivells d'LDL degut a que disminuiria l'efecte reductor d'LDL que provoca l'interferó.

Quin significat té que un canvi 3000 parell de bases cadena amunt d'un gen pugui modificar la resposta a un tractament antiviral? Ge *et al.* (2009) també observaren que aquest canvi presenta un gran desequilibri de lligament amb un canvi no sinònim situat en el gen IL28b (213A>G; K70R; rs8103142). Podria ser que aquest canvi alterés la funció del gen que codifica per IFN- λ 3 i explicaria les dades observades, però s'haurien de realitzar més estudis per definir completament el mecanisme biològic que actua (Thomas *et al.*, 2009). Altres estudis basats en el genotipatge de milers de polimorfismes comuns, donaren resultats semblants, posant de rellevància la regió IL28b (diferents polimorfismes amb desequilibris de lligament importants), com una de les més importants per determinar la resposta al tractament i l'aclariment espontani en pacients infectats amb el VHC (Rauch *et al.*, 2010a).

B. CTLA4 (rs5742909 i rs231775)

L'eliminació del virus, tant espontàniament com en resposta a la teràpia amb interferó i ribavirina, s'ha correlacionat amb una forta resposta immune que es presenta en forma d'una gran activació de les cèl·lules T (Vertuani *et al.*, 2002). L'antigen 4 dels limfòcits T citotòxics (CTLA4), correceptor codificat per un gen situat en el cromosoma 2 (2q33), s'expressa en les cèl·lules T CD4⁺ i CD8⁺ activades, produint una disminució de la funció de les cèl·lules T. Es trobaren dos SNPs situats en aquest gen que determinaven patrons de resposta diferents en pacients infectats amb VHC. Un dels SNP està situat en la posició -318 (C/T) (rs5742909) del gen, és a dir, en el promotor i l'altre en la posició +49 (A/G) (rs231775) produint el canvi d'aminoàcid alanina per treonina. Diferents estudis mostraren que l'al·lel +49G facilitaria l'eliminació del virus, tant de forma espontània com a resultat del tractament amb IFN i ribavirina (Yee *et al.*, 2003; Schott *et al.*, 2007). Schott *et al.* (2007) trobaren que aquest efecte només era significatiu en el cas dels homes, mentre que les dones no hi havia diferències significatives.

En quant a la combinació dels dos SNPs, es trobà que els pacients portadors de l'haplotip -318C/49G, i infectats amb el genotip 1 del virus, presentaven dues vegades més probabilitat de respondre positivament a la infecció que la resta de pacients. El per què

d'aquest efecte s'ha atribuït al fet que s'ha observat que l'haplotip complementari (-318T/+49A), provoca un augment de l'expressió de CTLA4, produint llavors una disminució de la resposta de les cèl·lules T (Yee *et al.*, 2003). Nischalke *et al.* (2010) obtingueren resultats semblants en pacients coinfectats amb VHC i VIH, observant que l'haplotip -318C/49G en homozigosi estava associat, de manera més significativa, a una major resposta viral sostinguda que els predictors clàssics de resposta (càrrega viral del VHC, edat, genotip del VHC, sexe).

C. LDLr (rs14158)

Els hepatòcits representen el principal lloc de replicació del VHC, encara que s'ha observat que també pot infectar altres tipus cel·lulars. Aquestes cèl·lules juguen un paper molt important en el metabolisme del colesterol en els mamífers i com s'ha esmentat abans, en la infecció pel VHC poden provocar un canvi en els paràmetres lipídics dels pacients. Els hepatòcits poden adquirir colesterol a dos nivells, el primer és mitjançant la seva síntesis i el segon captant el colesterol de les lipoproteïnes de baixa densitat (*low density lipoproteins* o LDLs) a través de l'endocitosi d'aquestes partícules utilitzant receptors específics, sent un d'ells l'LDLr (*low density lipoprotein receptor*). EL gen del receptor d'LDLs es localitza en el cromosoma 19 (19p13.2), està compost per 18 exons que determinen un transcrit de 5,3 Kb i codifiquen per un pèptid de 860 aminoàcids. Aquest receptor és una glicoproteïna de membrana, el lligand fisiològic del qual carrega amb el 65-70% del colesterol plasmàtic. Es va sospitar que el LDLr jugava un paper en la infecció pel VHC, donat la interacció d'aquests virus amb les lipoproteïnes. Diferents observacions fetes *in vitro* van en aquest sentit, com per exemple, que moduladors de l'expressió d'aquest receptor, també modulen en paral·lel la infecció pel VHC, o bé que les LDLs són una magnitud més eficient que les HDLs inhibint la infecció (Molina *et al.*, 2007; Ye, 2007). La disminució de l'expressió del receptor degut a alts nivells de colesterol podria interferir amb la propagació del VHC. De fet, elevats nivells de colesterol en sèrum són un factor predictiu d'una resposta viral sostinguda com a resultat de teràpies combinades amb interferó (Mas Marques *et al.*, 2009)

S'han dut a terme diferents estudis per analitzar si mutacions puntuals en el LDLr significaven algun canvi en l'evolució de la infecció pel VHC. Hennig *et al.* (2002) analitzaren diferents posicions del gen, trobant una relació entre els pacients que inicialment responien al tractament (amb interferó o combinat amb ribavirina) i canvis en una regió situada a l'extrem 3'UTR (anomenada també c.*52G>A o rs14158), malgrat alguns

d'aquests pacients temps endavant recaiguessin en la malaltia. Recentment s'han combinat els resultats obtinguts en aquest SNP amb els trobat a l'SNP rs12979860 del gen IL28b. Pareix ser que ambdós SNPs presenten un efecte sinèrgic en la resposta viral sostinguda (Pineda *et al.*, 2011; Rivero-Juárez *et al.*, 2012).

D. HFE

L'hemocromatosi primària o hemocromatosi hereditària és un trastorn genètic autosòmic recessiu del metabolisme del ferro. És la malaltia genètica més comú en poblacions d'origen caucàsic, i està caracteritzada per un increment de l'absorció de ferro a nivell de l'intestí i per la seva deposició en el fetge, pàncrees, cor i articulacions, entre altres òrgans. Les complicacions hepàtiques de l'hemocromatosi inclouen el desenvolupament de fibrosi, cirrosi i carcinoma hepatocel·lular (Njajou *et al.*, 2004).

El gen HFE està localitzat en el braç curt del cromosoma 6, concretament en la posició 6p21.3. Està constituït per 9,6 Kb organitzades en 7 exons i s'han pogut observar 11 transcrits de diferent longitud (Thenie *et al.*, 2000).

Les mutacions més freqüents del gen HFE són la C282Y i la H63D. La mutació C282Y està determinada per un canvi d'una G per una A en el nucleòtid 845 que determina el canvi d'una cisteïna per una tirosina, en l'aminoàcid 282 de la proteïna. La mutació H63D està determinada pel canvi d'una C per una G en el nucleòtid 187. Aquest canvi provoca que l'aminoàcid número 63 passi a ser un àcid aspàrtic en lloc d'una histidina (Hanson *et al.*, 2001). Un altre mutació relacionada amb el fenotip de la sobrecàrrega de ferro és la S65C, transversió d'una A per una T en el nucleòtid 193, que condueix a una substitució d'una serina per una cisteïna en l'aminoàcid 65 (Le Gac i Ferec, 2005).

La presència d'hemocromatosi es considera un factor de risc en el desenvolupament de la infecció per hepatitis C (Eisenbach *et al.*, 2004). Hi ha evidències que les mutacions C282Y i H63D en heterozigosi estan associades amb un elevat grau d'inflamació hepàtica, fibrosi més severa i una més freqüent progressió a cirrosi en pacients amb hepatitis C crònica (Geier *et al.*, 2004). Malgrat això, hi ha estudis que indiquen que seria possible que mutacions en aquest gen, podrien influenciar el desenvolupament o progressió d'altres malalties hepàtiques independentment dels efectes de la sobrecàrrega de ferro *per se*, i no sempre com a factors de risc. Això podria ser degut a que el producte del gen HFE és una proteïna amb una estructura homòloga a la del complex major d'histocompatibilitat de tipus

I i, per tant, podria afectar la resposta immunitària (Bonkovsky *et al.*, 2002). Bonkovsky *et al.* (2006) estudiaren la relació de les tres principals mutacions del gen HFE i la resposta viral sostinguda en pacients infectats amb el VHC i tractats amb interferó i ribavirina. Els resultats indicaren que les mutacions en el gen HFE, estaven associades amb una major resposta viral sostinguda, especialment evident en el cas de la mutació H63D. Carneiro *et al.* (2010) i Ishizu *et al.* (2012) donaren resultats en la mateixa direcció. S'ha de destacar que Espanya és el país del món amb una freqüència més elevada de la mutació H63D (Guix *et al.*, 2002).

E. Haplogrups mitocondrials

Els mitocondris són orgànuls essencials per la producció d'energia i el control de l'apoptosi en les cèl·lules. La fosforilació oxidativa duta a terme pel mitocondri, produeix ATP, mantén la temperatura corporal i genera espècies reactives d'oxigen (ROS). El DNA mitocondrial (mtDNA) presenta una elevada taxa de mutació, el que provoca una gran quantitat de polimorfismes. S'han descrit nombrosos conjunts de polimorfismes (o haplotips) que s'han agrupat en haplogrups. Aquests impliquen l'existència d'ancestres comú via materna (Wallace, 2005).

Les interaccions entre les infeccions víriques i la situació energètica de la cèl·lula suggeririen que variacions en el mtDNA podrien jugar un paper en la progressió de la malaltia produïda per la infecció. Els mitocondris són essencials per l'apoptosi, un procés clau del sistema immunitari dels hostes per respondre a una infecció viral. Es va observar que mutacions en algunes parts del mtDNA provocaven malalties que tenien uns efectes semblants als efectes produïts pel tractament del VIH. S'intentà correlacionar els haplogrups mitocondrials, amb la progressió de la infecció i s'observà que la progressió cap a la SIDA està associada amb una disminució de la síntesi de mtDNA. Els resultats també senyalaren que els haplogrups J i U5 presentaven una progressió accelerada cap a la SIDA, mentre que l'haplotip H presentava un risc elevat de lipoatròfia. Els resultats observats podrien tenir dues causes, la primera que degut a la forta estructura filogeogràfica dels haplogrups, les associacions observades estassin realment correlacionades amb unes propietats genètiques nuclears subjacents, que fossin diferents en els haplogrups estudiats. La segona raó es basa en el fet que els haplogrups relacionats amb una progressió accelerada de la malaltia estan relacionats també amb una menor producció d'ATP i ROS. (Hendrickson *et al.*, 2008; 2009).

En estudis realitzats en pacients coinfectats amb el VHC i el VIH, s'observà una associació entre els haplogrups HV i H i una menor taxa de progressió de fibrosi, mentre que l'haplogrup U s'associà a una progressió més ràpida de fibrosi. Per una altra banda l'haplogrup J fou relacionat amb un major risc aterogènic (García-Álvarez *et al.*, 2011; Micheloud *et al.*, 2011)



OBJECTIUS

El virus de l'hepatitis C és la causa més freqüent a nivell mundial de fibrosi hepàtica, cirrosi i desenvolupament d'un hepatocarcinoma. Una part dels pacients infectats, estan al mateix temps coinfectats amb el virus de la immunodeficiència humana, fet que modifica el curs de la malaltia. La determinació de les variables demogràfiques, genètiques i clíniques que puguin predir el desenvolupament d'una coinfecció VHC-VIH representa un gran repte. Els factors candidats a estar involucrats en aquest desenvolupament inclouen tant factors propis del virus com variables del propi hoste. Una de les estratègies per esbrinar els factors que puguin estar involucrats en el desenvolupament de la malaltia és la realització d'estudis retrospectius a partir de cohorts de pacients coinfectats.

La relació entre la variabilitat en la seqüència del virus de l'hepatitis C i el desenvolupament de la malaltia hepàtica no és simple i directe ja que, probablement, és de tipus multifactorial. En la primera part d'aquest estudi s'ha dut a terme una estimació dels factors associats amb el dany hepàtic emprant la metodologia GLM. Els objectius concrets són:

1. Estimar mitjançant índexs indirectes de caire no invasiu el dany hepàtic en pacients coinfectats amb el virus de l'hepatitis C i el virus de la immunodeficiència humana.
2. Correlacionar diferents dades demogràfiques, clíniques i les característiques epidemiològiques dels pacients coinfectats amb el VHC i el VIH, i la variabilitat genètica de la seqüència de la regió *core* i la regió *NS5B* del VHC amb el grau de dany hepàtic.
3. Aplicar i valorar un mètode estadístic basat en els "models lineals generalitzats (GLM)" per tal d'identificar les variables associades amb la progressió de la malaltia en el fetge.

El desenvolupament de la malaltia, a més, està influïda pel fet que els pacients estan sotmesos a tractament per alleujar els efectes de la infecció. En la segona part de l'estudi, s'han correlacionat dades pròpies de l'hoste i del virus amb la resposta o no al tractament per la infecció al VHC. Els objectius concrets són:

1. Determinar les freqüències gèniques i genotípiques de diferents polimorfismes genètics relacionats amb el desenvolupament de la malaltia en la població resident a Mallorca, tant coinfectada amb els virus VHC-VIH, com en una població control.

2. Relacionar característiques pròpies de l'hoste, incloent polimorfismes genètics, i del virus amb la resposta al tractament al virus VHC i amb el dany hepàtic a pacients coinfectats amb els virus VHC-VIH.

1. ESTUDI DEL GENOMA VÍRIC *versus* DANY HEPÀTIC

A. Pacients

Es varen seleccionar 55 pacients coinfectats amb els virus VHC i VIH procedents de l'hospital Son Llàtzer de Palma, Illes Balears (Espanya). L'estudi fou aprovat pel comitè institucional de revisió de protocols, i els pacients donaren el seu consentiment informat per emprar les seves mostres biològiques en aquesta recerca. Els criteris d'inclusió consistiren en què els pacients fossin seropositius i amb virèmia positiva pels dos virus (VHC i VIH), alhora que seronegatius pel virus de l'hepatitis B (VHB). S'utilitzà sèrum d'aquests pacients, que es va conservar a -70°C fins la seva utilització. Els 55 pacients foren inclosos en l'estudi de la regió *NS5B*, mentre que 37 foren inclosos en l'estudi de la regió *core*.

B. Variables

Es va construir una base de dades amb les diferents variables de cada pacient, les quals foren obtingudes revisant els historials clínics de forma retrospectiva. S'incloueren dades corresponents al sexe, edat, factors de risc per les infeccions d'ambdós virus (incloent la data de la primera exposició, si era possible), l'estat de la infecció per VIH (*1993 revised CDC classification*), administració de TARGA, consum d'alcohol (l'abús d'alcohol ha estat determinat per un consum mitjà diari de més de 50g durant més de dos anys), el genotip del VHC i diferents proves de laboratori rutinàries incloent el recompte CD4+ i les càrregues virals d'ambdós virus. Per tal de simplificar les anàlisis, l'estadi de la infecció per VIH va ser dividit en dues categories C3 i no C3, degut a que els estadis que no eren C3 tenien una freqüència molt baixa i, per tant es varen agrupar. La càrrega viral del VIH fou considerada com a baixa (B) quan era menor de 1000 còpies/ml i com a alta (A) quan era més gran de 1000 còpies/ml. Els marcadors bioquímics varen incloure els triglicèrids (TGC), el colesterol total (COL), el recompte plaquetari (PLA), la transaminasa glutamicopirúvica (GPT), la transaminasa glutamicooxalacètica (GOT) i la gammaglutamiltranspeptidasa (GGT), entre altres. El genotipatge del virus de l'hepatitis C es va dur a terme emprant el sistema *Versant HCV Genotype 2.0 (LIPA)* (Siemens). La quantificació de l'RNA del VHC i del VIH, es va dur a terme emprant els tests *AmpliPrep/COBAS Taqman HIV-1* i *AmpliPrep/CobasTaqman HCV*, ambdós de la casa Roche. El recompte de cèl·lules T CD4+

es va realitzar mitjançant una citometria de flux estàndard i es varen obtenir els resultats absoluts i relatius.

C. Mesura de la fibrosi hepàtica

El grau de dany hepàtic es va calcular mitjançant l'índex de Forns, l'APRI i el FIB-4 (Forns *et al.*, 2002; Sterling *et al.*, 2006; Vallet-Pichard *et al.*, 2007), degut al caràcter retrospectiu de l'estudi i a l'absència de biòpsies en molts dels casos. Els tres índexs foren utilitzats de forma individual com a variables de resposta en els models estadístics, però també es varen utilitzar simultàniament per veure correlacions entre els diferents índexs.

Aquest índexs es calculen utilitzant

L'índex de Forns (**IF**) (Forns *et al.*, 2002) es calculà segons la següent fórmula:

$$IF = 7,811 - 3,131 \cdot \ln(\text{recompte plaquetari}) + 0,781 \cdot \ln(GGT) + 3,467 \cdot \ln(\text{edat}) - 0,014 \cdot (\text{colesterol})$$

L'índex **APRI** (en anglès, *AST to platelet ratio index*) va ser calculat segons la fórmula desenvolupada per Wai *et al.* (2003):

$$APRI = \frac{\text{nivell GOT (U/ml)/LSN (U/ml)}}{\text{Recompte plaquetari (10}^9\text{/ml)}} \cdot 100$$

Sent LSN el límit superior normal de la GOT, en aquest cas s'ha considerat 40U/ml.

L'índex **FIB-4** (Sterling *et al.*, 2006) s'ha calculat segons l'equació següent:

$$FIB-4 = \text{edat(anys)} \cdot \frac{GOT(U/ml)}{\text{Recompte plaquetari (10}^9\text{/ml)} \cdot GPT(U/ml)^{1/2}}$$

En una altra fase de l'estudi, els índexs es varen corregir pel *temps estimat de la primera exposició*, per poder obtenir una aproximació de la taxa de desenvolupament de la fibrosi.

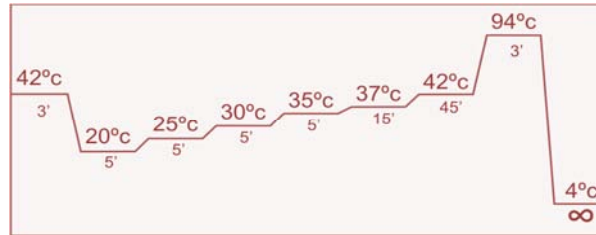
D. Seqüenciació i filogènia de les regions *core* i *NS5B*

L'RNA total del virus fou extret a partir de 400µl de sèrum sanguini, conservat a -70°C, emprant el *HighPure Viral RNA Kit* (Roche, Barcelona, Espanya).

A partir de l'RNA total es va realitzar una transcripció inversa emprant hexanucleòtids (*Random primers*, Promega, Madison, WI) com encebadors, per tal d'obtenir

el DNA complementari (cDNA) total. En un volum de 20µl es disposaren 10µl de l'RNA extret, 0,5mM de cada nucleòtid, 0,5µg d'hexanucleòtids, 20 unitats d'un enzim inhibidor d'RNAses (*RNasin*, Promega, Madison, WI), 100 unitats de la transcriptasa inversa (*M-MLV*, Promega, Madison, WI) i el tampó 5x pertinent.

La reacció de la retrotranscripció fou realitzada en un termociclador en les següents condicions:



Una vegada obtingut el cDNA total es va realitzar la PCR primària en un volum total de 25µl, per tal d'amplificar les 395pb de la regió *core* i les 331pb de la regió *NS5B* (nucleòtids 342-736 i 8274-8604, respectivament, de la seqüència de referència AF009606). El mètode a seguir fou el descrit per Ohno *et al.* (1997) amb modificacions. La reacció es va fer a partir de 2,5µl del cDNA obtingut i amb una concentració final de 0,2mM de cada nucleòtid, 0,4mM de cada encebador, 0,3 unitats de l'enzim *Taq polimerasa* (*Biotools*, Madrid, Espanya) i la quantitat corresponent del tampó 10x amb MgCl₂ adequat per l'enzim.

Per tal de poder amplificar tots els genotips en una única PCR, es va haver de realitzar una mescla d'encebadors, els quals, degut a la gran variabilitat dels virus, presentaven varies degeneracions. Per tant, com a encebadors genòmics (els que tenen el mateix sentit que l'RNA original del virus) es va utilitzar una mescla d'encebadors (Taula 2). En el cas de l'amplificació de la regió *core* es mesclaren el Core 1_1g i el Core 2_1g. En el cas de la regió *NS5B* l'NS5B 1_1g i l'NS5B 2_1g. En el cas dels encebadors antigenòmics (els complementaris a l'RNA original del virus), la mescla emprada consistia, en el cas de l'amplificació de la regió *core*, en els encebadors Core 1_1a i Core 2_1a. I en el cas de la regió *NS5B* en l'NS5B 1_1a i l'NS5B 2_1a.

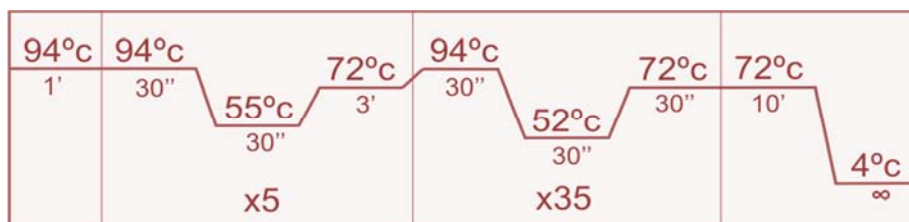
En ocasions, la PCR primària no fou suficient per amplificar adequadament el reduït nombre de còpies existents en el material genètic de partida. Per tal cosa, es realitzà una amplificació secundària amb les mateixes proporcions que la PCR primària, però emprant com a DNA motlle 2,5µl del producte de l'amplificació primària i uns encebadors més interns. Els nous encebadors emprats per l'amplificació de la regió *core* foren els

encebadors genòmics Core 1_2g i Core 2_2g. En el cas de la regió NS5B, els nous encebadors foren antigenòmics, l' NS5B 1_2, l' NS5B 2_2a i l' NS5B 3_2a.

Taula 2.- Encebadors utilitzats en les amplificacions primàries i secundàries de les regions *core* i *NS5B* del VHC.

Encebador	Seqüència 5'→3'
Core 1_1g	GGGAGGTCTCGTAGACCGTGCAYCATG
Core 2_1g	GGGAGGTCTCGTAGACCGTGCAACATG
Ns5b 1_1g	TATGATACYCGCTGYT [*] TYGACTC
Ns5b 2_1g	TATGAYACYCGCTGYT [*] TYGACTC
Core 1_1a	GAGMGGDATRTACCCCATGAGRTC GGC
Core 2_1a	GAGMGGDATRTATCCCATGAGRTC GGC
Ns5b 1_1a	GTACCTRGTCATAGCCTCCGTGAA
Ns5b 2_1a	TAY CTC GTCATRGCTCCGTGAA
Core 1_2g	AGACCGTGCAYCATGAGCAC
Core 2_2g	AGACCGTGCAACATGAGCAC
Ns5b 1_2a	CTCKYAGGCTCGCYGCRTCCCTC
Ns5b 2_2a	CTCKYAGGG YTGYYC YRTCCTC
Ns5b 3_2a	CTCSGAGGGCTCGKT [*] TRTCCTC

Les condicions emprades en ambdues reaccions foren les següents:



Per tal de comprovar que l'amplificació s'hagués produït de forma adequada (395pb) i suficient per la posterior seqüenciació, es realitzà una electroforesi carregant 8µl de les amplificacions en gels d'agarosa a l'1% (Figura 9).

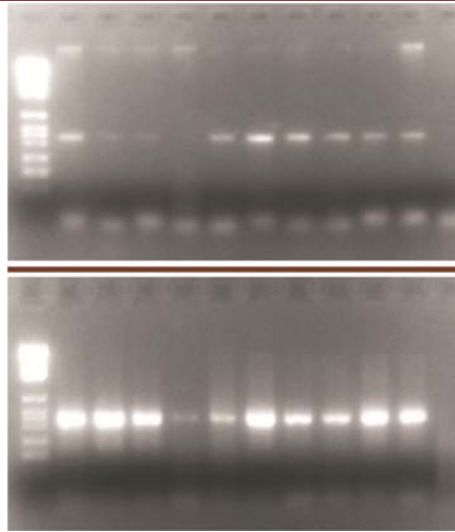


Figura 9. Electroforesi en gel d'agarosa a l'1% de l'amplificació primària (dalt) i secundària (baix) de la regió *core*. En el primer pouet es pot veure el marcador de pes molecular VI de Roche.

La purificació del producte de PCR es va realitzar emprant el kit de purificació *MSB Spin PCRapace* (Invitek, Berlin, Alemanya). La seqüenciació es va dur a terme mitjançant el *BigDye® Terminator v.3.1. cycle Sequencing Kit* (AppliedBiosystems, Foster City, CA) i, posteriorment, el producte d'aquesta reacció es va precipitar amb etanol i acetat sòdic.

Les seqüències es varen obtenir en un analitzador genètic 3130 (AppliedBiosystems, Foster City, CA), per després ser revisades i alineades de forma manual les dues cadenes, emprant el programa *Bioedit v.7.0.9.0* (Hall, 1999). L'assignació dels subtipus fou confirmada realitzant una cerca de tipus *Blast* en la base de dades *Los Alamos HCV sequence and immunology database* (Kuiken *et al.*, 2008).

El programa *MEGA 4* (Tamura *et al.*, 2007) es va emprar per construir els arbres filogenètics en els que es basaren les anàlisis. S'emprà el mètode de la Parsimònia màxima i un *bootstrap* de 2000 replicacions. S'utilitzà aquest mètode perquè s'ha mostrat que és més robust que el mètode de la Màxima versemblança, en el cas de que existeixin taxes de mutació heterogènies (Kolaczowski i Thornton, 2004). Els arbres finals es varen obtenir utilitzant l'arbre consens amb un 50% i un 75% de suport. La diversitat nucleotídica i haplotípica de la regió analitzada es va calcular usant el programa *DnaSP v5* (Librado i Rozas, 2009).

E. Anàlisis estadístiques

De totes les variables disponibles en la base de dades, es varen incloure unes o unes altres segons les diferents anàlisis realitzades. Per una banda, algunes variables van haver de ser eliminades per l'absència dels seus valors en alguns dels pacients estudiats, aquest fou el cas de la càrrega viral del VHC. En diferents casos, les variables relacionades més directament amb el càlcul de les variables de predicció del dany hepàtic, no varen ser incloses en el modelatge per tal d'evitar redundàncies.

Per una altra banda, es necessitava un mètode objectiu per poder incloure la informació genètica en els models estadístiques, l'elegit fou semblant a l'aplicat per Szymaragd *et al.* (2006). A cada clade dels dos arbres se'ls hi donà el valor d'una nova variable. Finalment, s'assignà a cada nova variable un valor binari (0 o 1) segons si el pacient estava inclòs en el clade corresponent a la variable o no.

Es va aplicar un mètode estadístic dins el marc dels mètodes lineals generalitzats (*Generalized Linear Methods* o GLM) (Venables i Ripley, 1999) emprant el programari lliure *The R project for Statistical Computing* (www.r-project.org). El "Model Lineal Generalitzat" és un desenvolupament de models lineals per tal de poder ajustar tant distribucions de resposta no-normal com transformacions a la linealitat, d'una manera clara i directa.

Un model lineal generalitzat es pot descriure segons les següents premisses:

i. Hi ha una variable de resposta d'interès (y) i una sèrie de variables de predicció (x_1, x_2, \dots), els valors de les quals influeixen en la distribució de la variable de resposta.

ii. Les variables de predicció influeixen la distribució de y sols a través d'una *única funció lineal*. Aquesta funció lineal s'anomena *predictor lineal* (*linear predictor*) i se sol expressar com:

$$y = \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_p \cdot x_p$$

La classe de models lineals generalitzats manejats pel paquet R, inclou distribucions de resposta (famílies) del tipus Gaussià, Binomial, Poisson, Gaussià Inversa, Gamma; i també models de Quasi-versemblança, on la distribució de resposta no està explícitament definida. Cada distribució de resposta admet una varietat de funcions d'enllaç (*link functions*) per connectar la mitjana amb el predictor lineal. La combinació de la funció d'enllaç

d'**identitat** ("*identity link*") amb la família Gaussiana produeix el **model lineal normal**, amb una funció d'enllaç del tipus $\beta_x = \mu$, i una funció mitjana $\mu = \beta_x$. Els estimadors de la màxima versemblança per aquest model són les estimes clàssiques de **mínims quadrats**. Quan s'utilitzen diferents variables de resposta al mateix temps en les anàlisis, emprant les mateixes variables de predicció, la combinació de la família Gaussiana amb la funció d'enllaç d'**identitat**, dona lloc a una **anàlisi normal multivariable**.

Per avaluar el millor model, s'empra el Criteri d'Informació d'Akaike (AIC), elegint el valor més baix de tots els models provats. L'AIC es coneix en estadística com un *penalized log-likelihood*, es a dir, un índex que penalitza els models. D'aquesta manera, en un model on es pot obtenir la *log-likelihood*:

$$AIC = -2 \cdot \log\text{-likelihood} + 2 \cdot (p + 1),$$

on p és el número de paràmetres en el model i el valor 1 s'afegeix per la variància estimada. L'AIC és útil perquè penalitza paràmetres superflus en el model, afegint $2 \cdot (p + 1)$ a la desviació. Quan comparem dos models, quan més baix és l'AIC, millor s'ajusta el model. Aquesta és la base de la simplificació automàtica del model emprant la funció *Step* en el programa R (Crawley, 2007).

Finalment, les comparacions entre els diferents índexs i les altres variables foren fetes emprant el programa *SPSS per Windows* (Rel. 15.0.1 2006 Chicago: SPSS Inc.).

2. ESTUDI DELS POLIMORFISMES EN PACIENTS versus RESPOSTA AL TRACTAMENT

A. Pacients i mostres control

Per realitzar aquest estudi, es seleccionaren 63 pacients coinfectats amb els virus VHC i VIH procedents de l'hospital Son Llàtzer de Palma. De la mateixa manera que anteriorment, l'estudi fou aprovat pel comitè institucional de revisió de protocols, i els pacients donaren els seu consentiment informat per emprar les seves mostres biològiques en aquesta recerca.

Els criteris d'inclusió foren els mateixos que en l'estudi anterior, és a dir, els pacients havien de ser seropositius i amb virèmia positiva pels dos virus (VHC i VIH), alhora que seronegatius pel virus de l'hepatitis B (VHB). Es varen emprar tres escovillons per extreure

cèl·lules de la mucosa bucal dels pacients i posteriorment es deixaren eixugar i s'emmagatzemaren a temperatura ambient.

Per una altra banda, es varen seleccionar 60 mostres de DNA anònimes (30 d'homes i 30 de dones) pertanyents a població resident disponibles en el laboratori de genètica de la Universitat de les Illes Balears, per tal d'analitzar els polimorfismes que no s'havien estudiat mai en aquesta població, en una població control.

B. Variables

Es varen seleccionar una gran quantitat de variables, obtingudes a partir dels historials clínics dels pacients, que després poguessin ser correlacionades amb els resultats de les anàlisis genètiques.

Les variables seleccionades es poden dividir en característiques bàsiques dels pacients (edat, sexe, índex de massa corporal), variables bioquímiques (transaminases, colesterol), hemograma (hemoglobina, leucòcits, plaquetes), variables relacionades amb la infecció (genotip, càrrega viral del VHC, via de transmissió), variables relacionades amb el possible efecte de la malaltia (recompte de cèl·lules CD4, dany hepàtic mesurat en la biòpsia i utilitzant *Fibroscan*, presència d'hepatomegàlia, esplenomegàlia i/o esteatosis hepàtica) i, finalment, variables relacionades amb la resposta al tractament del VHC (resposta a les 4, 12, 24 setmanes, resposta viral sostinguda, aclariment espontani). Es determinà que un pacient presentava resposta viral sostinguda quan la càrrega viral era indetectable 24 setmanes després d'haver acabat el tractament. Per una altra banda, es classificà com a aclariment espontani, el fet que un pacient presentés càrrega viral negativa, però resultés seropositiu pel VHC, sense haver rebut cap tipus de tractament.

C. Detecció dels polimorfismes

Extracció del DNA

L'extracció del DNA a partir de les cèl·lules de mucosa bucal dels escovillons es va realitzar utilitzant el mètode tradicional de fenol-cloroform i la posterior precipitació utilitzant etanol i NaCl.

Detecció dels polimorfismes

En aquest estudi es varen analitzar set polimorfismes d'un sol nucleòtid (SNPs) en les mostres dels pacients, a més de caracteritzar l'haplogrup mitocondrial. En les mostres

control només es van analitzar els polimorfismes que no s'havien estudiat a la població mallorquina, l'IL28b, els dos polimorfismes del gen CTLA4, l'S65C del gen HFE i el polimorfisme del gen LDLr. Pel genotipatge, es van emprar dues tècniques diferents, els RFLPs i la seqüenciació. A la Taula 3, es poden consultar els polimorfismes analitzats en aquest estudi, juntament amb la tècnica utilitzada per caracteritzar-los. En el cas de l'haplogrup mitocondrial, primerament es detectaren els pacients amb haplogrup H utilitzant RFLPs i després es seqüenciaren les regions hipervariables I i II (HvrI i HvrII) per tal de classificar la resta d'haplogrups.

Taula 3. Relació dels polimorfismes analitzats, detallant el gen al que pertanyen, el canvi puntual que produeixen i la tècnica utilitzada per determinar-lo.

Gen	Abreviatura	Polimorfisme	Canvi	Tècnica
Interleucina 28b	IL28b	rs12979860	T / A	RFLPs
Antigen 4 del limfòcit T citotòxic	CTLA4	+49 (rs231775)	G / A	RFLPs
	CTLA4	-318 (rs5742909)	C / T	RFLPs
Gen de l'hemocromatosi hereditària	HFE	C282Y	G / A	RFLPs
	HFE	H63D	C / G	RFLPs
	HFE	S65C	T / A	RFLPs
Receptor de lipoproteïnes de baixa densitat	LDLr	rs14158	A / G	Seqüenciació
Haplogrup mitocondrial	Haplo H			RFLPs
	Reste d'haplogrups			seqüenciació

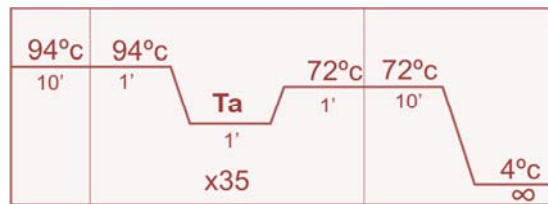
a. Encebadors

Per tal de la detectar aquests polimorfismes, primerament es va realitzar una amplificació emprant en cada cas els encebadors adients mostrats a la Taula 4.

b. Amplificacions

El primer pas per detectar els canvis, independentment de la tècnica seleccionada va ser una amplificació. A la Taula 5 es poden observar les condicions pel desenvolupament de les amplificacions.

Les condicions de les amplificacions dels polimorfismes nuclears es poden observar en la següent imatge, mentre que la temperatura d'anellament es pot consultar en la Taula 5.



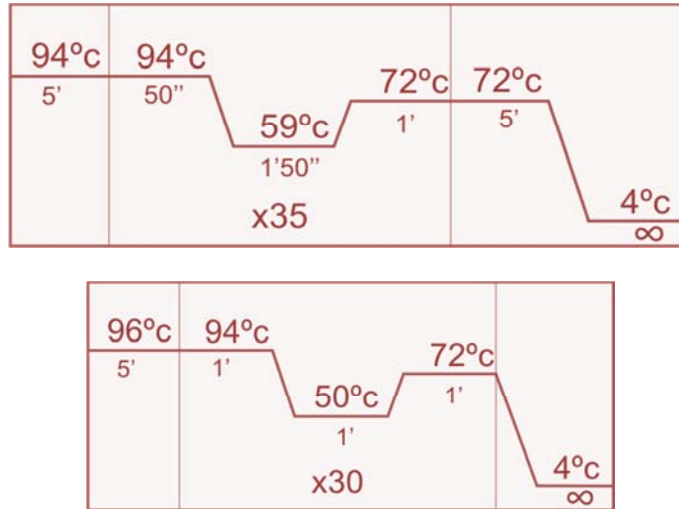
Taula 4. Encebadors utilitzats per la detecció dels diferents polimorfismes.

Encebador	Seqüència 5'→3'	Referències
IL28b F	GCTTATCGCATACGGCTAGG	Present treball
IL28b R	AGGGACCGCTACGTAAGTCA	Present treball
CTLA4 F	AAATGAATTGGACTGGATGGT	(Nischalke <i>et al.</i> , 2010)
CTLA4 R	AGCCAGCCAAGCCAGAT	
C282Y F	TGGCAAGGGTAAACAGATCC	(Feder <i>et al.</i> , 1996)
C282Y R	CTCAGGCACTCCTCTCAACC	
H63D/S65C F	ACATGGTTAAGGCCTGTTGC	
H63D/S65C R	GCCACATCTGGCTTGAAA TT	
LDLr F	TGGCAGAGACAGATGGTCAG	Present treball
LDLr R	CACTGTCCGAAGCCTGTTCT	Present treball
Haplo H L	CCGTAGGTGGCCTGACTGGC	(Picornell <i>et al.</i> , 2005)
Haplo H H	TGATGGCAAATACAGCTCCT	(Picornell <i>et al.</i> , 2005)
HVR I L	CTCCACCATTAGCACCCAAAG	(Vigilant <i>et al.</i> , 1989)
HVR I H	CACCATCCTCCGTGAAATCAATATCCC	
HVR II L	CTCACGGGAGCTCTCCATGC	
HVR II H	TGGCGGTATGCACTTTTAACA	

Taula 5. Reactius i condicions de la PCR utilitzades per realitzar les amplificacions emprades en aquest estudi.

Polimorfisme	DNA genòmic (ng)	Encebadors (cada un) (pM)	dNTPs (cada un) (mM)	Taq pol (U)	Temperatura d'anellament (°C)	Mida amplificat (pb)
IL28b	40	0,2	0,2	0,75	59	427
CTLA4 +49/-318	40	0,2	0,2	0,75	60	680
HFE C282Y	40	0,2	0,2	0,75	60	390
HFE H63D/S65C	40	0,2	0,2	0,75	60	208
LDLr	40	0,2	0,2	0,75	59	195
Haplogrup H	60	2,0	0,4	1,25	59	124
Seq Mitochondrial HVR I	80	0,2	0,4	1	50	453
Seq Mitochondrial HVR II	80	0,2	0,4	1	50	401

A continuació es representen les condicions de les amplificacions del DNA mitocondrial, primer l'amplificació per identificar els que pertanyen a l'haplotip H, i després l'amplificació per seqüenciar les regions hipervariables.



RFLPs: Digestió i gels d'agarosa

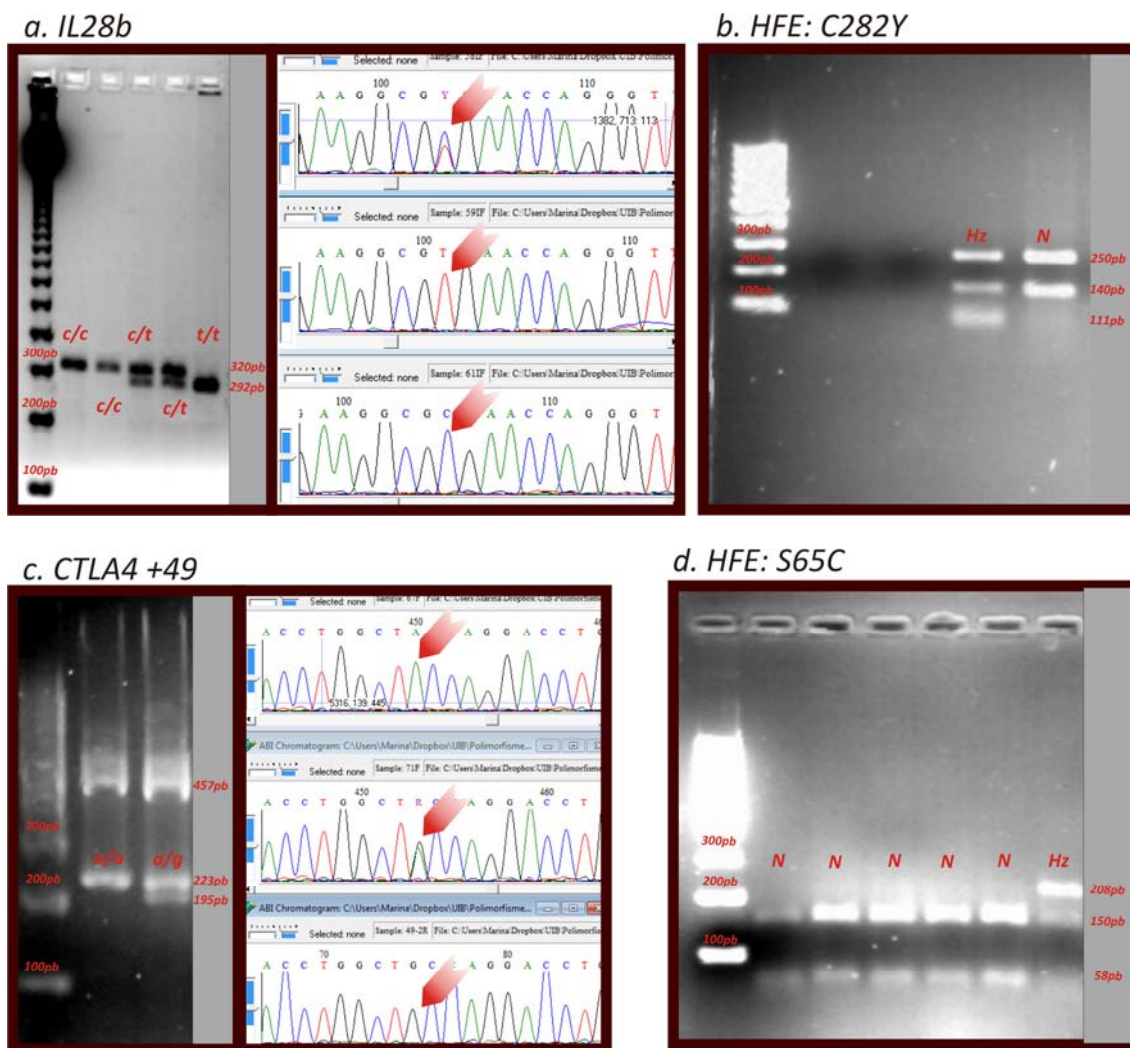
Tal com es pot veure a la Taula 3, es van dissenyar RFLPs (polimorfisme de longitud dels fragments de restricció) per determinar les variants dels gens IL28b, CTLA4, HFE i la determinació de l'haplogrup H. Una vegada obtingudes les amplificacions, tal com s'ha explicat a l'apartat b, es va procedir a realitzar les digestions amb els enzims de restricció específics. Per dissenyar les digestions no descrites a la bibliografia, es van utilitzar les eines de la pàgina web <http://tools.neb.com/NEBcutter2/>.

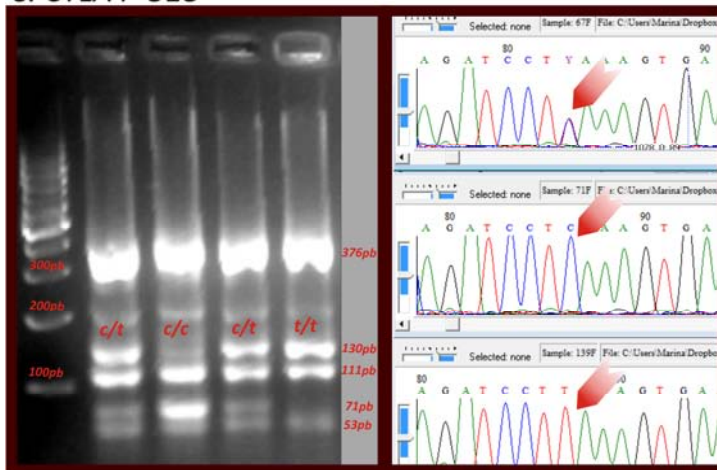
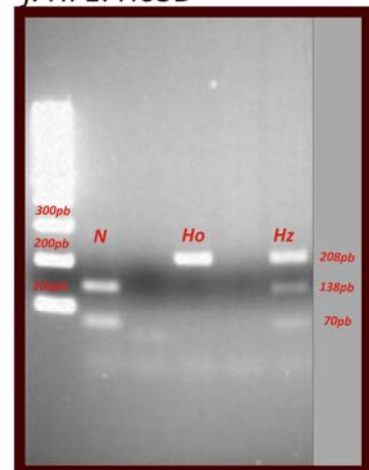
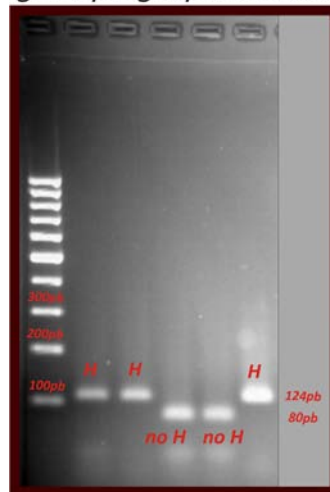
Taula 6. Enzims de restricció, temperatura de digestió i percentatge d'agarosa emprats en els RFLPs. S'indica la diana de restricció i el nucleòtid polimòrfic.

Polimorfisme	Enzim	Diana de l'enzim (nucleòtid polimòrfic)	Ta i temps	Percentatge gel agarosa	Referència
IL28b rs12979860	<i>Hpy166II</i>	GTN [^] NAC	37°C – 3h	2%	Present treball
CTLA4 +49	<i>Fnu4HI</i>	GC [^] NGC	37°C – 2h	4%	Present treball
CTLA4 -318	<i>MnI</i>	CCTC(N) ₇ [^]	37°C – 2h	4%	Present treball
HFE C282Y	<i>RsaI</i>	GT [^] AC	37°C – 3,5h	3%	(Lynas, 1997)
HFE H63D	<i>BcI</i>	T [^] GATCA	50°C – 2h	3%	(Lynas, 1997)
HFE S65C	<i>HinI</i>	G [^] ANTC	37°C – 3h	3%	Present treball
Haplogrup H	<i>AluI</i>	AG [^] CT	37°C – 3h	2%	(Picornell <i>et al.</i> , 2005)

A la Taula 6 es poden consultar els enzims de restricció específics utilitzats i la diana (detallant el punt de l'SNP).

Es varen realitzar electroforesis en gels d'agarosa mitjà per tal de separar els productes de digestió que es visualitzaren amb bromur d'etidi. A la Taula 6 s'especifica el percentatge d'agarosa utilitzat. Per comprovar la posada a punt dels nous mètodes, es van realitzar seqüenciacions de les regions a analitzar, seguint el mateix mètode explicat al punt següent. A continuació es mostren alguns exemples del patró de bandes obtinguts en cada cas, i, si escau, un petit fragment de la seqüenciació obtinguda.



e. *CTLA4 -318*f. *HFE: H63D*g. *Haplogrup mitocondrial H*

Seqüenciació

El polimorfisme situat en el gen *LDLr* i la determinació dels haplogrups mitocondrials distints a l'H, es van dur a terme emprant la seqüenciació. Una vegada realitzades les amplificacions, es procedí a la seva comprovació utilitzant gels d'agarosa al 1% emprant bromur d'etidi per la tinció de les bandes.

Després de la purificació dels productes amplificats amb el *kit Spin PCRapace kit* (Invitek, Berlin, Alemanya), es va dur a terme la reacció de seqüenciació amb el *BigDye® Terminator v.3.1. cycle Sequencing Kit* (Applied Biosystems, Foster City, CA), per ambdues cadenes, la precipitació amb etanol i l'anàlisi en un analitzador genètic automàtic ABI 3130 (Applied Biosystems, Foster City, CA).

Els fragments es visualitzaren mitjançant el programa *Bioedit Sequence Alignment Editor* (Hall, 1999). Per determinar l'haplogrup mitocondrial es va utilitzar el software online Haplogrep (Kloss-Brandstaeetter *et al.*, 2010).

D. Anàlisis estadístiques

Es calcularen les freqüències al·lèliques i genotípiques dels polimorfismes, així com l'equilibri de Hardy-Weinberg, en el grup de pacients i en el grup de controls. Per veure si existien diferències entre la distribució dels genotips entre pacients i controls es construïren taules de contingència i s'emprà l'estadístic Chi-quadrat o Fisher, segons el cas.

Per cada polimorfisme es calcularen diferents paràmetres per establir si existia alguna associació d'algun genotip amb alguna variable estadística. Per analitzar associacions amb variables qualitatives (sexe, genotip virus, via de transmissió, resultats de biòpsia, càrrega viral agrupada, hepatomegàlia, esplenomegàlia, esteatosi) es construïren taules de contingència i s'emprà l'estadístic Chi-quadrat o Fisher, segons el cas. Seguint el mateix procediment s'analitzaren els resultats de cada polimorfisme segons els paràmetres de resposta (resposta a les 4, 12 i 24 setmanes, resposta viral sostinguda, aclariment espontani).

Les variables quantitatives s'analitzaren segons seguien una distribució normal o no (dada obtinguda utilitzant el test de Kolmogorov-Smirnov). Les variables que seguien una distribució normal (índex de massa corporal, colesterol, plaquetes, albúmina, hemoglobina, leucòcits, FIB-4, recompte CD4 en percentatge, Fibroscan, bilirubina, edat) es varen analitzar usant la *t* d'Student o l'ANOVA, segons es comparessin dos o tres grups, respectivament.

Les variables quantitatives que no seguien una distribució normal (GOT, GPT, les càrregues virals, protrombina, APRI i el recompte de CD4 absolut) s'analitzaren usant les proves no paramètriques, Mann-Whitney o Kruskal-Wallis, segons es comparessin dos o tres grups, respectivament.

Els càlculs estadístics s'han realitzat utilitzant el paquet estadístic SPSS per *Windows* (Rel. 15.0.1 2006 Chicago: SPSS Inc).

0. ÍNDEX DE PUBLICACIONES

PUBLICACIÓ 1:

Relating the liver damage with hepatitis C virus polymorphism in core region and human variables in HIV-1-coinfected patients

Autors: Marina Matas, Antònia Picornell, Carmen Cifuentes, Antoni Payeras, Antoni Bassa, Francesc Homar, F. Xavier López-Labrador, Andrés Moya, Maria M. Ramon, José A. Castro

Infection, Genetics and Evolution (2010) 10: 1252-1261

PUBLICACIÓ 2:

Generalized Linear Model (GLM) framework for the association of host variables and viral strains with liver fibrosis in HCV/HIV coinfecting patients

Autors: Marina Matas, Antònia Picornell, Carmen Cifuentes, Antoni Payeras, Antoni Bassa, Francesc Homar, Fernando González-Candelas, F. Xavier López-Labrador, Andrés Moya, Maria M. Ramon, José A. Castro

Acceptat: Infection, Genetics and Evolution (Octubre 2012)

PUBLICACIÓ 3:

Relating the outcome of HCV infection with different host SNP polymorphisms in a Majorcan coinfecting HCV-HIV population treated with pegIFN-RIB

Autors: Marina Matas, Antònia Picornell, Carmen Cifuentes, Antoni Payeras, Antoni Bassa, Francesc Homar, Fernando González-Candelas, F. Xavier López-Labrador, Andrés Moya, Maria M. Ramon, José A. Castro

En preparació

Matas M, Picornell A, Cifuentes C, Payeras A, Bassa A, Homar F, López-Labrador FX, Moya A, Ramon MM, Castro JA (2010). Relating the liver damage with hepatitis C virus polymorphism in core region and human variables in HIV-1-coinfected patients. *Infect Genet Evol* 10, 1252-1261.

1. Generalized Linear Model (GLM) framework for the association of host variables and viral strains with liver fibrosis in HCV/HIV coinfecting patients.
-

Authors: Marina Matas ^a, Antònia Picornell ^a, Carmen Cifuentes ^b, Antoni Payeras ^b, Antoni Bassa ^b, Francesc Homar ^b, Fernando González-Candelas ^c, F. Xavier López-Labrador ^c, Andrés Moya ^c, Maria M. Ramon ^a, José A. Castro ^{a*}

^aInstitut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Departament de Biologia, Universitat de les Illes Balears, Campus de la UIB, Palma de Mallorca, Illes Balears, 07122, Spain

^bUnitat de Malalties Infeccioses, Hospital Son Llàtzer, Palma de Mallorca, Illes Balears, Spain

^cCentre Superior d'Investigació en Salut Pública (CSISP), Av. de Catalunya 21, 46020 València; Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, Apartado de Correos 22085, 46071 València, CIBER en Epidemiología y Salud Pública (CIBEResp), Instituto de Salud Carlos III, Spain

Running title: GLM in liver damage by HCV-HIV co-infection

A. Abstract

Chronic hepatitis C virus (HCV) infection is the main cause of advanced and end-stage liver disease world-wide, and an important factor of morbidity and mortality in Human Immunodeficiency virus-1 (HIV-1) co-infected individuals. Whereas the genetic variability of HCV has been studied extensively in monoinfected patients, comprehensive analyses of both patient and virus characteristics are still scarce in HCV/HIV co-infection. In order to find correlates for liver damage, we sought to analyze demographic, epidemiological and clinical features of HCV/HIV co-infected patients along with the genetic makeup of HCV (viral subtypes and lineage studied by nucleotide sequencing and phylogenetic analysis of the NS5B region). We used the generalized linear model (GLM) methodology in order to integrate data from the virus and the infected host to find predictors for liver damage. The degree of liver disease was evaluated indirectly by means of two indexes (APRI and FIB-4) and accounting for the time since infection, to estimate fibrosis progression rates. Our analyses identified a reduced number of variables (both from the virus and the host) implicated in liver damage, which included the stage of HIV infection, levels of gamma-glutamyl transferase and cholesterol, and some distinct HCV phylogenetic clades.

Keywords: hepatitis C virus, liver damage, HCV-HIV co-infection, Generalized linear models, GLM.

B. Introduction

The World Health Organization (WHO) estimates that about 3% of the world population is infected with the hepatitis C virus (HCV). This means that there are about 200 million carriers of the virus in the world (Thomson and Finch, 2005). The prevalence of HCV infection is highly variable, ranging from less than 1% in northern Europe to 15-20 % in Egypt (Alter, 2007). In recent decades, an increase in prevalence has been detected as a result of the virus spreading through networks of intravenous drug users (IDU) (Thomson and Finch, 2005). This common transmission route is responsible for the current high incidence of HCV/HIV co-infection in developed countries.

HCV is the only species of the *Hepacivirus* genus within the *Flaviviridae* family. The HCV genome consists of a single-stranded RNA molecule with positive polarity of about 9.6 Kb, encoding a polyprotein of 3,000 amino acids. This polyprotein is cleaved by proteases of the host and the virus itself in three structural and seven non-structural proteins, including an RNA-dependent RNA polymerase (NS5B). The NS5B protein is composed by 591 amino acids and is the catalytic element of the virus replication machinery. This protein has become a clear target for new therapies against HCV, although its genetic variability may limit the efficacy of specific inhibitors because of naturally-occurring resistant variants (Beaulieu, 2009; Gaudieri *et al.*, 2009; Shi *et al.*, 2008). The genetic variability of HCV was revealed in the early 90's, when the methodology to sequence and classify the different variants was developed. As a result, HCV has been classified into six well-defined groups or clusters, known as genotypes, which are in turn divided into about 80 subtypes (Simmonds *et al.*, 2005).

The relative frequencies of genotypes and subtypes in different populations are not uniform across risk groups, reflecting population changes, age of patients and routes of transmission. Therefore, the distribution of genotypes and subtypes varies along time and space. In industrialized countries, subtype 1a is usually found in young individuals with IDU as the main risk factor, while subtype 1b is more common in older people with a history of transfusions (Pybus *et al.*, 2005). In Spain, HCV prevalence is similar to that of other industrialized countries, being genotype 1 the most prevalent followed by genotype 3 (Bruguera and Forns, 2006; Esteban *et al.*, 2008), although the distribution of subtypes presents some differences compared to other European countries (López-Labrador *et al.*, 1997), with a higher proportion of subtype 1b infections. HCV NS5B gene has been widely

used for genotype/subtype/isolate classification of viral strains (Simmonds *et al.*, 2005). Although several studies have analyzed the intra-patient genetic variability of the virus and its relevance for disease progression (Farci and Purcell, 2000; Farci *et al.* 2000), the application of quasispecies analysis to clinical practise may be unpractical (Curran *et al.*, 2002; Sullivan *et al.*, 2007). Instead, a more practical approach may be trying to identify genetically diverse viral strains with potential differential effects on liver fibrosis (López-Labrador *et al.*, 2006).

The natural history of an HCV infection is difficult to elucidate. The time elapsed between the establishment of a chronic infection and the development of end stage liver disease ranges from 15 to 40 years. Another problem in the study of infection arises because chronic infection, often including the initiation of the cirrhosis process, occurs without symptoms (Missiha *et al.*, 2008). According to various studies (see Seef (2009) and references herein), the rate of spontaneous clearance of the virus is about 15-25% in infected adults. A higher rate is found in children and young women, reaching 45%. This group also has the lowest percentages in the development of cirrhosis and hepatocellular carcinoma. Although HCV is not considered to be cytophatic, the hallmark pathogenic effect of chronic infection is the generation of liver fibrosis with time (McCaughan and George, 2004). Fibrosis progression ultimately leads to cirrhosis and, consequently, to end-stage liver disease. Several studies have revealed differences in prognosis and disease development among the different HCV genotypes. For instance, Poynard *et al.* (2003) found that genotype 3 is possibly associated with further development of fibrosis (see Probst *et al.*, 2011, for a recent review).

HCV-HIV co-infection prevalence is estimated around 2% of the general population in Western countries, but in patients with a history of IDU the prevalence rises to 70-95% (Alter M.J., 2006). Co-infection is associated with a higher mortality rate than that produced by each virus separately (Rotman and Liang, 2009). Available data suggest that co-infected patients have an increased risk of liver-related disease and a faster evolution of the disease than HCV-monoinfected patients (Matthews and Dore, 2008). Overall, co-infected patients have a higher HCV viral load, more severe liver inflammation, an accelerated progression to cirrhosis, and higher mortality rates (Alter H., 2006). Co-infected patients should receive antiretroviral treatment for HIV, which is often associated with increased levels of transaminases, indicating an impairment of liver function probably contributing to increased fibrosis (Poynard *et al.*, 2003). Some studies indicate that a co-

infected patient is 2 to 10 times more likely to suffer liver toxicity when treated with highly active antiretroviral therapy (HAART) (Thomson and Finch, 2005).

Finding predictors for rapid fibrosis progression in co-infected patients may guide clinicians in the choice for the most appropriate treatment regime for each individual patient. In the present study we have searched for a potential association of HCV viral genetic factors (different viral strains distinguished by their NS5B gene sequence) and simple demographic, clinical, and epidemiological characteristics of HCV-HIV co-infected patients with the degree of liver damage estimated by using non-invasive fibrosis measures. We applied the GLM methodology (Venables and Ripley, 1999), previously used in the investigation of genetic variables of viruses involved in the outcome of chronic infection (Matas *et al.*, 2010; Szmargd *et al.*, 2006). This statistical methodology can relate different variables (responses and predictors) and select the most relevant ones, combining qualitative and quantitative variables in a wide and versatile way.

C. Materials and methods

Patients

Fifty-five HCV-HIV co-infected patients were chosen retrospectively from the Infectious Diseases Unit of Son Llàtzer Hospital, Palma, Illes Balears (Spain). Inclusion criteria were prior informed consent to use biological samples for medical research, documented seropositivity and available viremia levels for both HCV and HIV. No patient was co-infected with hepatitis B virus (HBV). Stored serum or plasma samples (-70°C) were used for the study. The study was approved by the local institutional ethics committee.

Variables

The following demographic, epidemiological, clinical, and laboratory variables were obtained by chart review and recorded in detail when possible: age, sex, risk factors for HIV and HCV infection (route of transmission, RofT), date of initial exposure (THCV), HIV disease staging (1993 Revised CDC Classification), highly-active antiretroviral therapy (HAART), alcohol consumption (alcohol abuse was defined by an average daily consumption >50g for >2 years), HCV genotype, routine laboratory tests including CD4-positive (CD4+), cell counts and HIV and HCV viral loads. Biochemical markers included: serum triglycerides; serum alanine-aminotransferase (ALT); and serum aspartate-aminotransferase (AST); gamma-glutamyl transpeptidase (GGT); cholesterol levels (COL)

among others. HCV genotyping was performed using the VERSANT HCV Genotype 2.0 (LIPA) (Siemens). HIV and HCV RNA quantification was performed using the COBAS AmpliPrep/COBAS TaqMan HIV-1 Test, versión 2.0 (Roche), and COBAS AmpliPrep/COBAS TaqMan HCV (Roche). CD4+ T cell counts were determined by standard flow cytometry and given in both absolute and relative numbers.

Liver damage

The degree of liver damage was measured by means of the indexes APRI and FIB-4 (Sterling *et al.*, 2006; Vallet-Pichard *et al.*, 2007), due to the retrospective nature of the study and lack of liver biopsies. The two indexes were used as response variables in the statistical modelling.

The **APRI** (AST to platelet ratio) index was calculated as described (Wai *et al.*, 2003) by using the formula:

$$\mathbf{APRI} = ((AST\ level\ (U/L) / ULN\ (U/L)) / (Platelet\ count\ (10^9/L))) \times 100,$$

being ULN the upper limit of normal AST level. We have considered this as 40.

The **FIB-4** index was calculated as described (Sterling *et al.*, 2006), using the formula:

$$\mathbf{FIB-4} = age\ (yr) \times AST\ (U/L) / (Platelet\ count\ (10^9/L) \times (ALT\ (U/L))^{1/2})$$

HCV sequencing and phylogeny of the NS5B region

Viral RNA was extracted from 400 ml of serum or plasma stored at -70°C with the High Pure Viral RNA kit (Roche Diagnostics, Barcelona, Spain). Reverse transcription and PCR amplification of 331 nucleotides from the NS5B region (nucleotides 8274 to 8604 in the reference sequence AF009606) were performed following the method described by Ohno *et al.* (1997) with little modifications. Amplified products were purified by means of column affinity with the MSB Spin PCRapace kit (Invitex, Berlin, Germany) and direct bidirectional sequencing was done with the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and analysed in a 3130 Genetic Analyzer (Applied Biosystems). All chromatograms were verified visually and both strands assembled using Bioedit v.7.0.9.0 (Hall, 1999). The deduced amino acid sequences were obtained with the same software according to the correct reading. HCV subtype assignment was confirmed by BLAST searches using the interface at the *Los Alamos HCV Sequence and Immunology Database* (Kuiken *et al.*, 2008).

Phylogenetic trees were constructed from the protein sequence using MEGA (Tamura *et al.*, 2007) with the *maximum parsimony* criterion and 2,000 bootstrap replicates to evaluate the support for tree nodes. We performed a close-neighbour interchange (CNI) with search level = 1. The initial tree for CNI searches was a random addition tree, with 10 replicates. We used the consensus tree, and only the nodes supported by at least the 50% majority rule were considered as significant. We used *maximum parsimony* following the criterion found in the bibliography (Kolaczowski and Thornton, 2004; Matas *et al.*, 2010; Szymaragd *et al.*, 2006). Nevertheless, we also performed the analysis with *maximum likelihood* obtaining similar results.

Statistical methodology

We applied a GLM framework (Venables and Ripley, 1999) using the free software R (<http://www.r-project.org>). We followed the methodology explained in detail in Szymaragd *et al.* (2006) and Matas *et al.* (2010). In summary, generalized linear modelling is a development of linear models to accommodate both non-normal response distributions and transformations to linearity in a clean and straightforward way. A generalized linear model may be described in terms of the following sequential assumptions: (i) there is a response of interest, y , and (ii) predictor variables x_1, x_2, \dots , whose values influence the distribution of y through a *single linear function*. This linear function is called the linear predictor, and is usually written as $y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$

The maximum-likelihood estimates for this model are the ordinary least-squares estimates. When several response variables are calculated and used simultaneously in the analysis with the same predictor variables, the combination of a Gaussian family with "identity" links gives a multivariate normal analysis.

The Akaike's information criterion (*AIC*) was used to evaluate the best of the different models (the one with the lowest *AIC* value). The *AIC* value is computed as a penalized log-likelihood with $AIC = -2 \times \log\text{-likelihood} + 2(p + 1)$, where p is the number of parameters in the model, and 1 is added for the estimated variance. *AIC* is useful because it explicitly penalizes any superfluous parameters in the model, by adding $2(p + 1)$ to the deviance. When comparing two or more models, the smaller the *AIC* the better the fit to the observed data. Finally, comparisons between different markers and other variables were done using *SPSS* for Windows, Rel. 15.0.1. 2006. Chicago: SPSS Inc. Data was expressed as absolute values for categorical variables, and means and standard deviations for continuous variables.

D. Results

Baseline characteristics of study subjects

The main features of the studied cohort (n=55) are shown in Table 1, including demographic and viral characteristics of the HIV-HCV co-infected subjects and the variables considered as fibrosis predictors for the total cohort. Sex was not distributed evenly as male patients more than doubled female patients. Intravenous drug use was the main risk factor for HCV/HIV co-infection, and most patients had low HIV viral load (below 400 IU/mL), reported no alcohol consumption, and were also receiving HAART. The individual details of the studied cohort can be found in Appendix 1. The correlation between the two indirect fibrosis indexes was very high ($r=0.95$), not only because we used the relation AST/Platelet in both expressions, but also because both formula try to be accurate predictors of liver damage.

The distribution of HCV subtypes found in the present cohort is consistent with that previously described by Cifuentes *et al.* (2004) in Majorcan HCV-HIV co-infected population. Hence, this cohort can be considered as representative of the population of co-infected patients in our region.

HCV NS5B phylogeny

The *maximum parsimony* consensus phylogenetic tree, based on the amino acid sequence of NS5B, grouped the sequences in clades. For the GLM analysis, only the clades with a bootstrap support >50% were selected and they were identified by capital letters in alphabetical order (A-K). In the choice of the nodes we followed strictly the clusters showed by the tree. In this way, the nodes are nested, in the sense that clade B also includes clade A; the clade D includes the clades A, B and C; and so on (Appendix 1). These eleven clades were introduced in the dataset to be included in the statistical modelling as predictive variables. In Figure 1 is also indicated the isolates classified in viral subtypes (clades I-V). The aim of these two classifications was evaluating the relevance of NS5B diversity *versus* viral subtypes in the progression of fibrosis. Samples 25LN-2a and 17LN-3a from two patients were excluded from this analysis because their fibrosis indexes were considered as outliers. In fact, when the analyses were carried out including these outlier samples, a trend selecting the patient with the subtype 2a was detected. This effect was not so strong for the sample 17LN-3a, probably because it was not the only representative of this subtype.

GLM and other Statistical analyses

We performed two types of analyses under the GLM framework. First, we included the information on the clades identified in the parsimony phylogenetic tree, and later we included only the subtype data. In both cases, this new information was coded as a binary variable = 1 or 0, when the sample belonged or not to a clade or genotype, respectively (Matas *et al.*, 2010; Szmaragd *et al.*, 2006). To avoid redundant information, the relationship of genotypes (GEN in Appendix 1) was not used in the analysis. Appendix 1 shows the new binary variables for the dataset. The following predictor variables were considered in the analyses: sex, alcohol consumption (ENO), route of transmission (RoT), date of initial exposure (THCV), highly-active antiretroviral therapy (HAART), CD4-positive (CD4+), HIV disease staging (SHIV), HIV viral load (LHIV), serum triglycerides (TGC), gamma-glutamyl transpeptidase (GGT), cholesterol levels (COL), along with the phylogenetic cluster variables (A-K or I-V) to model the degree of fibrosis (FIB4 and APRI as response variables). The variables used to calculate APRI and FIB4 indexes (AST, ALT, AGE and Platelet) were not included in the model to avoid redundancies and inconsistencies in the analyses. At the same time, the variables that were not available in many patients in the cohort were not introduced in the model (as HCV viral load). The models calculated by GLM, considered the two fibrosis indexes at once and used jointly as the response variable. Both indexes followed a normal distribution and, therefore, the response fitted a bivariate normal distribution (Gaussian family with identity link). The results of the GLM analyses for phylogenetic and genotype clades are shown in Table 2. In the analyses with phylogenetic clusters the best fit included GGT levels with a positive coefficient (Table 2a). The Non-C3 HIV stage and cholesterol level showed a negative association with liver damage. Phylogenetic cluster F, which included all subtype 3a isolates in the parsimony tree, was significantly associated with the degree of fibrosis, while cluster E (with two 3a isolates) showed a negative association. In the analyses with subtypes (Table 2b) (Roman numerals in Figure 1 and Appendix 1), only GGT levels were statistically positively correlated with fibrosis as in the previous analysis. Nevertheless, in this analysis no subtype variable was selected.

We also performed the same analyses including a correction for the estimated time from infection (FIB4c and APRIc in Table 2c). This correction was made by dividing these two parameters by the corresponding estimated time from infection (THCV), which was then removed as predictor variable to avoid redundancies. Analyses were also conducted

with the phylogenetic clusters and with viral subtypes. In both cases, the results were identical: GGT levels were positively related to fibrosis while cholesterol levels presented a negative relationship. No clades or subtypes were selected in any case.

E. Discussion

We carried out an analysis of factors potentially involved in the degree of liver fibrosis in patients with HCV-HIV co-infection, analyzing several patient characteristics and HCV NS5B nucleotide sequences within the GLM methodology (Matas *et al.*, 2010; Szmaragd *et al.*, 2006). This methodology allows the combination of different clinical, demographic and genetic factors and the use of quantitative and qualitative variables in a single analysis. The results show a significantly reduced model with only few significant predictors. The analyses showed the association between the response variable (the degree of liver fibrosis), GGT and cholesterol levels. While particular HCV subtype 3a NS5B phylogenetic clades were associated with fibrosis, they were not associated with fibrosis progression through time. A potential limitation of our study replies in its retrospective nature because the number of biopsies available was low, and therefore the information regarding real fibrosis was poor. We thus estimated the level of fibrosis, using two widely-used indirect indexes: APRI and FIB-4 (Sterling *et al.*, 2006; Vallet-Pichard *et al.*, 2007; Wai *et al.*, 2003). It is well known that these indexes may be inaccurate, although they seem much more accurate for HCV-HIV co-infected patients (Ramos Paesa *et al.*, 2007), despite the fact that platelet and transaminase levels could be altered by the HIV itself, or by HAART.

Accordingly, the results reported here should be interpreted with caution. Some authors have raised the question of whether liver biopsies have real diagnostic value for the evaluation of fibrosis, because of sampling error (Bedossa *et al.*, 2003); although, nowadays, the detection of the fibrosis by biopsy is still considered as the *gold standard*. Besides, the invasiveness of the procedure and the possible severe complications makes desirable finding other surrogate variables evaluating fibrosis, or using other non-invasive techniques, such as liver elastography.

In order to increase the accuracy of the indirect measures of fibrosis, we used the two indexes simultaneously in the analyses as a single multivariate response variable. Our analyses were Gaussian instead of multinomial (as indicated by Szmaragd *et al.*, 2006), because the fibrosis indexes represented a broad range of values. For each model, the host

variables and phylogenetic or subtype HCV clades (A-K or I-V) were used as potential predictors of the degree of fibrosis in our cohort of HCV–HIV co-infected patients.

We used a *maximum parsimony* approach in the construction of phylogeny tree of HCV NS5B amino acid sequences. The most remarkable feature of the tree is that in some genotypes, the samples were grouped into subclades, thus increasing the resolution of the subsequent fibrosis analysis. For the viral subtype data analysis, we have followed the same methodology as in the case of the clades. That is, each subtype was recoded into a new binary variable in order to systematize the method. To avoid interference with the final model, we excluded outlier isolates for the GLM analyses (i.e. the 25LN-2a isolate which was the only representative of subtype 2a in our cohort), minimizing bias towards outliers, as detected in previous analyses (Matas *et al.*, 2010).

In the analyses based on phylogenetic clusters and baseline measures of fibrosis (Table 2a), GGT was the only significant variable positively associated with fibrosis while the non-C3 HIV stage and cholesterol level (COL) were also significantly associated, but in a negative way. These results are in accordance with clinical finding from other studies. First, Kau *et al.* (2008) found elevated levels of GGT in patients with hepatic steatosis and advanced fibrosis. Second, the classification of HIV infection in a C3 stage is an AIDS-indicator condition, together with extremely low CD4+ T-cell counts. Some studies suggest that an advanced stage of HIV infection leads to an increased risk of liver disease in HCV-HIV co-infected patients, so an earlier stage of HIV infection (non-C3 stage) would be consistent with lower fibrosis levels (Matthews and Dore, 2008). Finally, a recent study (Piver *et al.*, 2010) has shown that lipid metabolism is greatly altered in the course of HCV infection, showing that levels of triglycerides and cholesterol in patients who have developed chronic hepatitis C (and therefore likely to develop more fibrosis), are lower than those of patients who have spontaneously cleared the virus. This may explain the negative relationship of cholesterol levels to liver damage. However, most studies in this field have been performed in cohorts of mono-infected patients, and the findings of the present study are now showing these relationships also in co-infected HCV/HIV individuals.

Finding genetic characteristics of HCV strains potentially associated with fibrosis may be clinically-relevant. In the current study, one NS5B clade (clade F, that includes all the 3a subtypes) was positively associated by GLM as a predictor variable, while another (clade E, that only grouped two 3a subtype patients, the 3LN-3a and 12LN-3a) was

negatively associated; showing, in this way, that the intragenetic variability detected in the 3a subtype is important in these two patients and, also, different in relation with the other patients with the same viral subtype. No NS5B clade was selected in the analyses with subtype data only, even when we used fibrosis indexes adjusted by the estimated time from infection. These results indicate that the consideration of intra-genotypic variability allows making associations beyond the viral genotype or subtype, thus increasing the resolution or specificity of the analysis. It is noteworthy that the HCV NS5B genetic diversity was highest in isolates from subtype 3a, which included a subclade (E) highly differentiated in the tree from the rest of isolates from this subtype genotype at the amino acid level. The relevance of the GLM method is patent when comparing the analysis with the viral subtypes *versus* the phylogenetic tree. Whereas in the first case (Table 2b) GGT was the only variable selected (and any subtype), using phylogeny, as discussed above, some subclades from the tree were also selected as predictors, emphasizing the relevance of viral diversity analysis based on nucleotidic or aminoacidic sequences and not only based on viral subtype determination.

We also performed the GLM analyses using fibrosis indexes corrected by the estimated time from infection to somehow account for the rate of progression of fibrosis. In these analyses the same results were obtained using phylogenetic clusters or subtype data. In any case, no clade or subtype was selected for inclusion in the model while GGT was selected with a positive coefficient and cholesterol with a negative one. Thus, these results may indicate that these two variables are the best independent surrogates for the development of hepatic fibrosis in our cohort. However, it should be noted that, whereas the rate of fibrosis progression might be more relevant than a single measure of fibrosis, the indexes used in the analyses were described as single measures of fibrosis. Therefore, the correction by the time from infection could in fact distort the indexes, and results obtained using this correction estimate should be interpreted with caution.

Among the potential explanations for the differential level of fibrosis between patients infected with different NS5B clades of subtype 3a, one could speculate that there may be some differences at the protein level that could determine these changes. Further analyzing of the sequences indicates that both clades differ in one amino acid position (pos. 273 in NS5B and 2793 in the polyprotein). In the samples belonging to clade E, this position is occupied by a proline, while in the other 3a isolates this position is occupied by a glutamine, but the potential effect of this variation in viral replication is unknown.

F. Conclusions

The GLM methodology allows integrations of data from both the virus and the host to find predictors for liver damage. The stage of HIV infection, levels of gamma-glutamyl transferase and cholesterol were identified as potential surrogates for liver fibrosis in HCV-HIV co-infected individuals. The role of some distinct HCV NS5B phylogenetic clades in liver fibrosis needs to be further evaluated in HCV-HIV co-infection.

G. Acknowledgements

This work was supported by grants 4326/2007 and 6557/2010 from the Direcció General de Recerca, Desenvolupament Tecnològic i Innovació. Conselleria d'Economia, Hisenda i Innovació, Govern Balear, Spain; projects PI10/00512, PI10/01734, and CIBEResp, Instituto de Salud Carlos III, Spain; projects BFU2008 03000BMC and BFU2011-24112 from Ministerio de Ciencia e Innovación, Spain. F.X.L. holds a P.I. position supported by the Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Spanish Ministry of Science.

H. References

- 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm Rep.* 41, 1–19.
- Alter, H., 2006. Viral hepatitis. *Hepatology* 43, S230–234.
- Alter, M.J., 2006. Epidemiology of viral hepatitis and HIV co-infection. *J. Hepatol.* 44, S6–9.
- Alter, M.J., 2007. Epidemiology of hepatitis C virus infection. *World J Gastroenterol.* 13, 2436–2441.
- Beaulieu, P.L., 2009. Recent advances in the development of NS5B polymerase inhibitors for the treatment of hepatitis C virus infection. *Expert Opin. Ther. Pat.* 19, 145–164.
- Bedossa, P., Dargere, D., Paradis, V., 2003. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 38, 1449–1457.
- Bruguera, M., Forns, X., 2006. Hepatitis C in Spain. *Med. Clin. (Barc.)* 127, 113–117.
- Cifuentes, C., Juan, M., Sastre, M., Payeras, A., Fernández-Baca, V., Homar, F., Gallegos, C., and Bassa, A., 2004. Coinfected by HIV-HCV: Epidemiological data. First

- Genotype determination in a sample from Mallorca Island. 27th World congress of Internal Medicine. Granada, Spain.
- Curran, R., Jameson, C.L., Craggs, J.K., Grabowska, A.M., Thomson, B.J., Robins, A., Irving, W.L., Ball, J.K., 2002. Evolutionary trends of the first hypervariable region of the hepatitis C virus E2 protein in individuals with differing liver disease severity. *J. Gen. Virol.* 83, 11–23.
- Esteban, J.I., Sauleda, S., Quer, J., 2008. The changing epidemiology of hepatitis C virus infection in Europe. *J. Hepatol.* 48, 148–162.
- Farci, P., Purcell, R.H. 2000. Clinical significance of hepatitis C virus genotypes and quasispecies. *Sem. Liver Dis.* 20, 103–122.
- Farci, P., Shimoda, A., Coiana, A., Diaz, G., Peddis, G., Melpolder, J.C., Strazzer, A., Chien, D.Y., Munoz, S.J., Balestrieri, A., Purcell, R.H., Alter, H.J. 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science* 288, 339–344.
- Gaudieri, S., Rauch, A., Pfafferott, K., Barnes, E., Cheng, W., McCaughan, G., Shackel, N., Jeffrey, G.P., Mollison, L., Baker, R., Furrer, H., Günthard, H.F., Freitas, E., Humphreys, I., Klenerman, P., Mallal, S., James, I., Roberts, S., Nolan, D., Lucas, M., 2009. Hepatitis C virus drug resistance and immune-driven adaptations: relevance to new antiviral therapy. *Hepatology* 49, 1069–1082.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95–98.
- Kau, A., Vermehren, J., Sarrazin, C., 2008. Treatment predictors of a sustained virologic response in hepatitis B and C. *J. Hepatol.* 49, 634–651.
- Kolaczowski, B., Thornton, J.W., 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* 431, 980–984.
- Kuiken, C., Hraber, P., Thurmond, J., Yusim, K., 2008. The hepatitis C sequence database in Los Alamos. *Nucleic Acids Res.* 36, D512–516.
- López-Labrador, F.X., Ampurdanés, S., Forns, X., Castells, A., Sáiz, J.C., Costa, J., Bruix, J., Sánchez Tapias, J.M., Jiménez de Anta, M.T., Rodés, J., 1997. Hepatitis C virus (HCV) genotypes in Spanish patients with HCV infection: relationship between HCV genotype 1b, cirrhosis and hepatocellular carcinoma. *J. Hepatol.* 27, 959–965.
- López-Labrador, F.X., Bracho, A., Berenguer, M., Coscolla, M., Rayón, J.M., Prieto, M., Carrasco, D., Gómez, M.D., Moya, A., González-Candelas, F., 2006. Genetic similarity of hepatitis C virus and fibrosis progression in chronic and recurrent infection after liver transplantation. *J. Viral Hep.* 13, 104–115.
- Matas, M., Picornell, A., Cifuentes, C., Payeras, A., Bassa, A., Homar, F., López-Labrador, F.X., Moya, A., Ramon, M.M., Castro, J.A., 2010. Relating the liver damage with hepatitis C virus polymorphism in core region and human variables in HIV-1-coinfected patients. *Infect. Genet. Evol.* 10, 1252–1261.

- Matthews, G.V., Dore, G.J., 2008. HIV and hepatitis C coinfection. *J. Gastroenterol. and Hepatol.* 23, 1000–1008.
- Missiha, S.B., Ostrowski, M., Heathcote, E.J., 2008. Disease progression in chronic hepatitis C: modifiable and nonmodifiable factors. *Gastroenterology* 134, 1699–1714.
- McCaughan, G.W., George, J., 2004. Fibrosis progression in chronic hepatitis C virus infection. *Gut* 53, 318–321.
- Ohno, O., Mizokami, M., Wu, R.R., Saleh, M.G., Ohba, K., Orito, E., Mukaide, M., Williams, R., Lau, J.Y., 1997. New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. *J. Clin. Microbiol.* 35, 201–207.
- Piver, E., Roingeard, P., Pages, J.C., 2010. The cell biology of hepatitis C virus (HCV) lipid addiction: molecular mechanisms and its potential importance in the clinic. *Int. J. Biochem. Cell Biol.* 42, 869–879.
- Poynard, T., Yuen, M.F., Ratziu, V., Lai, C.L., 2003. Viral hepatitis C. *Lancet* 362, 2095–2100.
- Probst, A., Dang, T., Bochud, M., Egger, M., Negro, F., Bochud, P.Y., 2011. Role of Hepatitis C virus genotype 3 in liver fibrosis progression – a systematic review and meta-analysis. *J. Vir. Hep.* 18, 745–759.
- Pybus, O.G., Cochrane, A., Holmes, E.C., Simmonds, P., 2005. The hepatitis C virus epidemic among injecting drug users. *Infect. Genet. Evol.* 5, 131–139.
- Ramos Paesa, C., Marcilla, F., López, G., Hueso, E., Pascual, A., Aguirre, J.M., 2007. Valuation of APRI and Forns models for non-invasive diagnosis of fibrosis in patients with hepatitis C in coinfecting and non-coinfecting with HIV. *An. Med. Interna (Madrid)* 24, 369–374.
- Rotman, Y., Liang, T.J., 2009. Coinfection with hepatitis C virus and human immunodeficiency virus: virological, immunological, and clinical outcomes. *J. Virol.* 83, 7366–7374.
- Seeff, L.B., 2009. The history of the "natural history" of hepatitis C (1968-2009). *Liver Int.* 29, 89–99.
- Shi, S.T., Herlihy, K.J., Graham, J.P., Fuhrman, S.A., Doan, C., Parge, H., Hickey, M., Gao, J., Yu, X., Chau, F., Gonzalez, J., Li, H., Lewis, C., Patick, A.K., Duggal, R., 2008. In vitro resistance study of AG-021541, a novel nonnucleoside inhibitor of the hepatitis C virus RNA-dependent RNA polymerase. *Antimicrob. Agents Chemother.* 52, 675–683.
- Simmonds, P., Bukh, J., Combet, C., Deleage, G., Enomoto, N., Feinstone, S., Halfon, P., Inchauspe, G., Kuiken, C., Maertens, G., Mizokami, M., Murphy, D.G., Okamoto, H., Pawlotsky, J.M., Penin, F., Sablon, E., Shin-I, T., Stuyver, L.J., Thiel, H.J., Viazov, S., Weiner, A.J., Widell, A., 2005. Consensus proposals for a unified system

- of nomenclature of hepatitis C virus genotypes. *Hepatology* 42, 962–973.
- Sterling, R.K., Lissen, E., Clumeck, N., Sola, R., Correa, M.C., Montaner, J., Sulkowski, M., Torriani, F.J., Dieterich, D.T., Thomas, D.L., Messinger, D., Nelson, M., 2006. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43, 1317–1325.
- Sullivan, D.G., Bruden, D., Deubner, H., McArdle, S., Chung, M., Christensen, C., Hennessy, T., Homan, C., Williams, J., McMahon, B.J., Gretch, D.R., 2007. Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. *J. Infect. Dis.* 196, 239–248.
- Szmaragd, C., Nichols, R.A., Balloux, F., 2006. A novel approach to characterise pathogen candidate genetic polymorphisms involved in clinical outcome. *Infect. Genet. Evol.* 6, 38–45.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596–1599.
- Thomson, B.J., Finch, R.G., 2005. Hepatitis C virus infection. *Clin. Microbiol. Infect.* 11, 86–94.
- Vallet-Pichard, A., Mallet, V., Nalpas, B., Verkarre, V., Nalpas, A., Dhalluin-Venier, V., Fontaine, H., Pol, S., 2007. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology* 46, 32–36.
- Venables, W.N. and Ripley, B.D., 1999. Generalized linear models. *Modern Applied Statistics with S*. Springer, New York, pp.183–210.
- Wai, C.T., Greenson, J.K., Fontana, R.J., Kalbfleisch, J.D., Marrero, J.A., Conjeevaram, H.S., Lok, A.S., 2003. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 38, 518–526.

I. Figures and Tables

Table 1. Baseline demographic, epidemiological and clinical characteristics of HIV–HCV-co-infected subjects (n=55) and variables considered as liver damage predictors. Number of patients or mean and standard deviation (SD) are indicated in the variables.

HCV SUBTYPES		1a=27 1b=6 2a=1 3a=10 4a=3 4d=8			
Demographic and epidemiological characteristics of subjects			Clinical characteristics of subjects		
			Mean	SD	
SEX	16 females	39 males	CD4 cell count, Cell/mm ³ (CD4)	388.67	212.96
AGE, years	Mean=39.82	SD=6.85	Time since HCV infection, years (THCV)	15.25	6.14
Route of Transmisión (Roft)	Heterosexual contact (HET)=4	Intravenous drug use (IDU)=51	GGT level, U/mL	103.31	96.33
Alcohol Consumption (ENO)	No=43	Yes=12	AST level, U/mL	55.09	33.02
Highly active antiretroviral therapy (HAART)	No=8	Yes=47	ALT level, U/mL	60.84	49.20
HIV stage (SHIV)	C3stage (C3)=23	Non C3 (NC3)=32	TGC level, mg/dL	173.75	157.65
HIV viral load (LHIV)	High (H)=12	Low (L)=43	COL total level, mg/dL	165.02	62.56
			Platelet count, 10 ⁹ /L (PLA)	198.35	96.85
			FIB4	2.02	2.27
			APRI	0.96	1.02

GGT: gamma-glutamyltransferase; AST: aspartate transaminase; ALT: alanine transaminase; TGC: triglycerides; COL: cholesterol. Alcohol consumption Yes: >50g for >2 years. HIV stage according 1993 revised CDC classification. HIV viral load High: >400 IU/mL

Table 2. GLM analyses. (a) Analysis is based on the phylogenetic clustering (cut-off BS=50%, letters A-K in the tree); (b) Analysis is based on viral subtypes (roman numerals I-IV in the tree); (c) Analysis with the indexes corrected by the time of infection (THCV, not included in the analyses); both analyses (with phylogenetic clusters and subtypes) gave the same result.

(a)

glm(formula = FIB4 + APRI ~ SHIV + GGT + COL + F + E, family = gaussian, data = NS5B) **AIC: 185.9**

Coefficients	Estimate	Std.Error	t value	Pr(> t)
(Intercept)	3.162	0.648	4.877	0.000***
SHIVNC3	-0.941	0.386	-2.435	0.019*
GGT	0.007	0.002	3.298	0.002**
COL	-0.006	0.003	-2.133	0.038*
Clade F	1.624	0.551	2.948	0.005**
Clade E	-2.990	1.069	-2.796	0.007**

Null deviance: 132.58 on 52 df, residual deviance: 79.46 on 47 df

(b)

glm(formula = FIB4 + APRI ~ GGT, family = gaussian, data = NS5B) **AIC: 195.79**

Coefficients	Estimate	Std.Error	t value	Pr(> t)
(Intercept)	1.690	0.323	5.233	0.000***
GGT	0.007	0.002	3.111	0.003**

Null deviance: 132.58 on 52 df, residual deviance: 111.43 on 51 df

(c)

glm(formula = FIB4c + APRIc ~ GGT + COL, family = gaussian, data = NS5B) **AIC: -61.00**

Coefficients	Estimate	Std.Error	t value	Pr(> t)
(Intercept)	0.267	0.055	4.847	0.000***
GGT	0.0005	0.0002	2.603	0.012*
COL	-0.0008	0.0003	-2.857	0.006**

Null deviance: 1.09 on 52 df, residual deviance: 0.84 on 50 df

* p<0.05; ** p<0.01; *** p<0.001

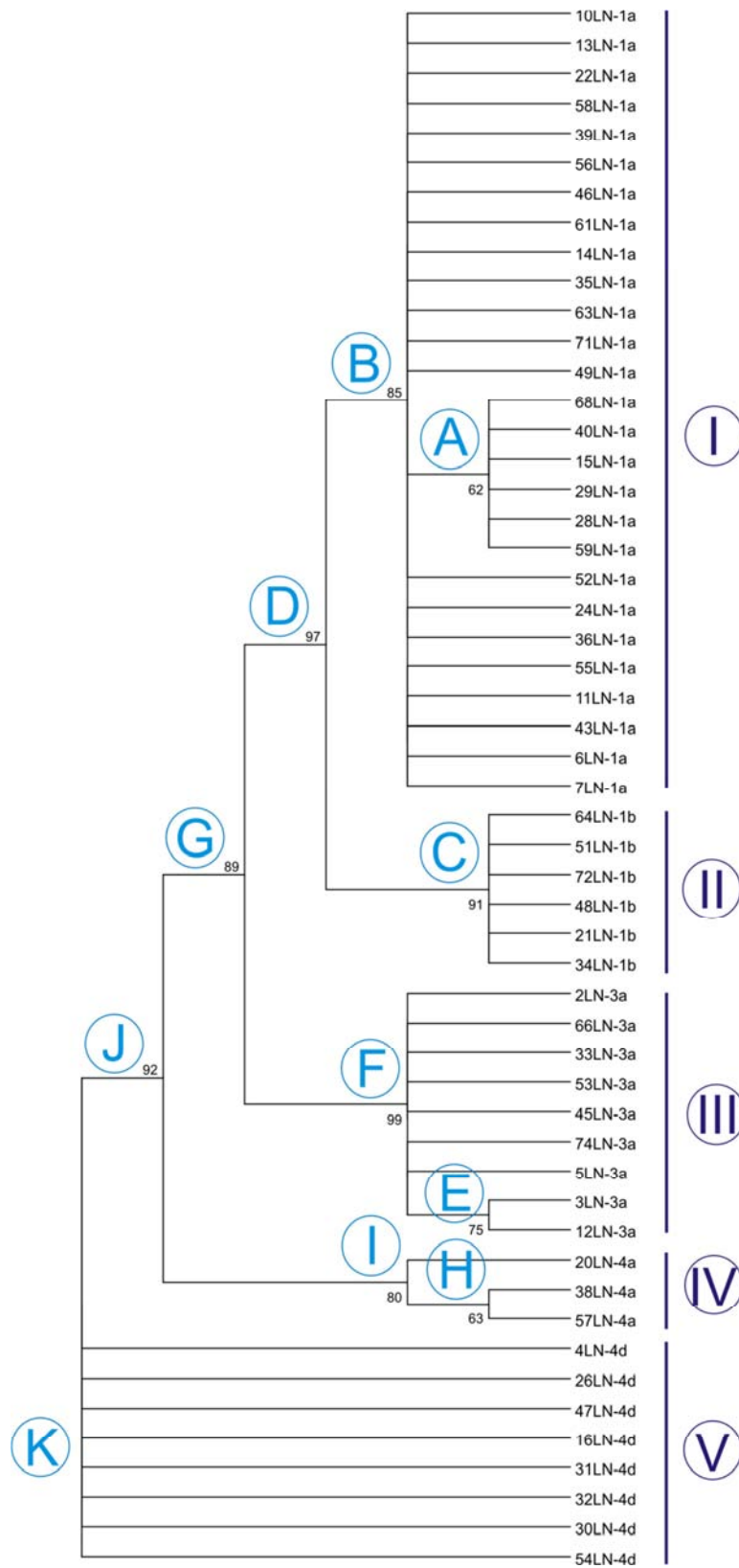


Fig. 1. The maximum parsimony consensus tree for the NS5B region, supported by bootstrap (2000 replicates). Letters A–K represent clades with >50% support, and roman numerals I–V the subtypes.

Patients	50% Majority Rule Variables													Subtype Variables													Fibrosis Indexes												
	GEN	A	B	C	D	E	F	G	H	I	J	K	I	II	III	IV	V	AGE	SEX	RcH	ENO	SHIV	CD4	LHV	HAART	AST	ALT	GGT	PLA	COL	TGC	THCV	FIB4	APRI	FIBac	APRlc			
68LNta	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	44	F	IDU	N	C3	507	L	Y	42	78	46	136	129	70	15	1.53	0.77	0.102	0.051			
40LNta	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	38	M	IDU	Y	NC3	844	L	Y	37	35	91	123	185	148	18	1.93	0.75	0.107	0.042			
15LNta	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	32	M	IDU	Y	NC3	489	L	Y	94	97	312	256	131	61	6	1.19	0.92	0.199	0.153			
29LNta	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	40	M	IDU	Y	NC3	176	H	Y	49	62	188	177	221	259	19	1.41	0.89	0.074	0.036			
28LNta	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	46	F	IDU	N	NC3	562	H	N	78	80	218	138	124	15	2.40	1.17	0.160	0.078				
59LNta	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	56	M	IDU	N	C3	273	L	Y	37	46	73	196	218	260	15	1.56	0.47	0.104	0.031			
10LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	42	F	IDU	N	C3	64	L	Y	43	72	22	171	170	56	15	1.24	0.63	0.083	0.042			
13LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	39	F	IDU	Y	C3	520	L	Y	94	51	128	180	189	177	17	2.85	1.31	0.168	0.027			
22LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	37	M	IDU	N	NC3	345	L	Y	31	66	141	215	181	193	17	0.66	0.36	0.039	0.021			
58LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	40	M	IDU	N	NC3	575	L	Y	35	39	74	289	171	76	8	0.78	0.30	0.097	0.038			
39LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	35	M	IDU	N	C3	4	H	Y	105	51	176	105	86	276	20	4.90	2.50	0.245	0.125			
58LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	35	F	IDU	N	C3	316	L	Y	65	74	134	187	199	284	18	1.41	0.87	0.079	0.048			
46LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	37	F	HET	N	NC3	430	L	Y	39	59	20	323	209	38	9	0.58	0.30	0.065	0.034			
61LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	31	M	IDU	N	NC3	147	L	Y	57	31	87	78	89	242	12	4.07	1.83	0.339	0.152			
14LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	30	F	IDU	N	NC3	480	L	Y	62	73	80	190	166	61	21	1.15	0.82	0.055	0.039			
35LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	35	F	IDU	Y	NC3	813	H	N	21	20	81	282	57	110	2	0.63	0.20	0.314	0.100			
63LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	50	M	IDU	N	C3	497	L	Y	34	40	78	293	236	295	7	0.92	0.29	0.131	0.041			
71LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	40	M	IDU	N	NC3	403	L	Y	58	65	93	130	218	240	18	2.21	1.12	0.123	0.062			
49LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	44	M	IDU	N	NC3	486	L	Y	37	37	73	253	150	141	8	1.06	0.37	0.132	0.046			
52LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	50	M	IDU	N	NC3	441	L	Y	62	84	159	314	267	715	18	1.08	0.49	0.060	0.027			
24LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	48	M	IDU	N	C3	300	L	Y	34	30	22	178	177	153	17	1.67	0.48	0.098	0.028			
36LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	39	M	IDU	N	NC3	468	H	N	50	47	147	214	124	76	21	1.33	0.58	0.063	0.028			
55LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	42	F	IDU	N	C3	335	L	Y	111	142	42	188	235	258	13	2.08	1.48	0.160	0.114			
11LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	34	F	IDU	Y	C3	239	L	Y	18	13	22	149	445	1029	16	1.14	0.30	0.071	0.019			
43LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	30	M	IDU	Y	C3	340	H	Y	86	81	166	194	116	200	14	1.48	1.11	0.105	0.079			
06LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	43	M	IDU	N	C3	197	L	Y	49	42	21	132	138	153	22	2.46	0.93	0.112	0.042			
07LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	37	M	IDU	N	C3	66	L	Y	35	38	145	208	298	194	21	1.00	0.42	0.047	0.020			
64LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	40	M	IDU	N	NC3	326	L	N	21	17	100	640	141	270	24	0.32	0.08	0.013	0.003			
51LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	37	M	IDU	N	C3	637	L	Y	32	34	44	193	148	85	21	1.05	0.41	0.050	0.020			
72LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	39	M	IDU	N	NC3	242	L	Y	46	93	185	181	166	66	21	1.03	0.64	0.049	0.030			
48LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	33	M	IDU	Y	NC3	393	H	Y	36	56	99	184	210	159	11	0.86	0.49	0.078	0.044			
21LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	62	M	IDU	N	NC3	372	L	Y	35	34	218	215	181	193	15	1.73	0.41	0.115	0.027			
34LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	34	M	IDU	Y	NC3	794	H	N	70	55	81	92	115	57	15	3.49	1.90	0.233	0.127			
25LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	43	F	IDU	N	NC3	152	H	N	26	27	19	26	132	69	22	8.28	2.50	0.376	0.114			
17LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	45	M	HET	Y	NC3	191	L	Y	84	63	95	31	99	72	3	15.36	6.77	5.121	2.258			
03LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	41	M	IDU	N	NC3	145	H	N	14	12	24	394	115	154	2	0.42	0.09	0.210	0.044			
12LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	42	M	IDU	Y	C3	602	L	Y	20	17	49	267	150	170	15	0.76	0.19	0.051	0.012			
02LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	38	F	IDU	N	NC3	763	L	Y	121	222	41	244	75	21	12	1.26	1.24	0.105	0.103			
68LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	38	M	IDU	N	C3	400	L	Y	88	63	121	105	179	51	23	4.01	2.10	0.174	0.091			
33LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	31	F	IDU	N	NC3	465	L	Y	82	54	92	191	141	9	3.76	2.23	0.418	0.248				
53LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	40	M	IDU	Y	NC3	484	L	Y	45	70	50	175	110	160	24	1.23	0.64	0.051	0.027			
45LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	41	M	IDU	N	NC3	251	L	Y	126	74	72	182	120	79	21	3.30	1.73	0.157	0.082			
74LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	41	M	IDU	Y	NC3	356	L	Y	45	64	54	134	148	206	8	1.72	0.84	0.215	0.105			
05LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	40	M	IDU	N	NC3	136	H	Y	65	69	164	110	152	99	17	2.85	1.48	0.167	0.087			
38LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	31	M	IDU	Y	NC3	273	L	Y	34	31	31	185	108	72	10	1.02	0.46	0.102	0.046			
57LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	51	M	IDU	N	C3	370	L	Y	75	94	257	169	147	146	26	2.33	1.11	0.090	0.043			
20LNta	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0	0	42	M	IDU	N	NC3	310	L	Y	25	37	139	178	225	156	15	0.97	0.35	0.065	0.023			
04LNta	0	1	0	1	0	1	0	1	0																														

2. Relating the outcome of HCV infection with different host SNP polymorphisms in a Majorcan coinfecting HCV-HIV population treated with pegIFN-RIB.
-

Marina Matas ^a, Antònia Picornell ^a, Carmen Cifuentes ^b, Antoni Payeras ^b, Antoni Bassa ^b, Francesc Homar ^b, Fernando González-Candelas ^c, F. Xavier López-Labrador ^c, Andrés Moya ^c, Maria M. Ramon ^a, José A. Castro ^{a*}

^aInstitut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Departament de Biologia, Universitat de les Illes Balears, Campus de la UIB, Palma de Mallorca, Illes Balears, 07122, Spain

^bUnitat de Malalties Infeccioses, Hospital Son Llàtzer, Palma de Mallorca, Illes Balears, Spain

^cCentre Superior d'Investigació en Salut Pública (CSISP), Av. de Catalunya 21, 46020 València; Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, Apartado de Correos 22085, 46071 València, CIBER en Epidemiología y Salud Pública (CIBEResp), Instituto de Salud Carlos III, Spain

A. Abstract

Hepatitis C virus (HCV) is the main cause of acute and chronic liver diseases, which are manifested as chronic hepatitis, cirrhosis or hepatocellular carcinoma. In the present study, we relate different host SNPs in IL28b, CTLA4, LDLr, and HFE genes and mtDNA haplogroups with the outcome of HCV infection in a Mallorcan coinfecting HCV-HIV population treated with pegIFN-RBV. Sixty-three HCV-HIV coinfecting patients and 59 anonymous controls were selected. The population frequency of the SNP IL28b gene is similar to frequency found in the European American cohort. However, C allele frequency is lower than those found in other Spanish cohorts. The frequencies of CTLA4 and LDLr polymorphisms are similar to those found in other populations. Differences between the cohort of patients and the control population were only found in H63D mutation. The other polymorphisms and the mitochondrial haplogroups do not show any difference between patients and controls. In the cohort of patients, the most significant data are found in treatment response at 4 weeks, so it seems that the CC genotype (IL28b) promote a rapid response to treatment. No relationship between the gene LDLr and the response to treatment was found. The spontaneous clearance is higher in patients with G allele in CTLA4+49. Mutations in the HFE gene seem to have no connection with any variable studied. Only one IL28b CC genotype patient infected by genotype 3 virus has not achieved SVR, on the other hand, only one IL28b TT genotype patient infected by genotype 1 or 4 virus has achieved SVR. This study wanted to highlight the importance of knowing the genetic background of the patients before starting a treatment taking in account the specific characteristics of the cohort.

Keywords: HCV-HIV co-infection, mtDNA haplogroups, LDLr, CTLA4, IL28b, HFE

B. Introduction

Hepatitis C virus (HCV) is the main cause of acute and chronic liver diseases, which are manifested as chronic hepatitis, cirrhosis or hepatocellular carcinoma (Alter, 2007).

The World Health Organization (WHO) estimates that up to 3% of world population has been infected with hepatitis C; this means that there are about 170 million carriers of the virus (Thomson and Finch, 2005). The prevalence of HCV infection ranges from less than 1% in Northern Europe to more than 2% on average in North Africa. The highest HCV prevalence (around 15-20%) can be found in Egypt (Alter, 2007). Transmission of hepatitis C occurs primarily through blood or derivatives, being mainly due to blood transfusions and contaminated needles and syringe sharing. With the establishment of HCV blood screening (in most countries since 1991), HCV-related cases of transfusion has almost disappeared. HCV prevalence has increased in recent decades as a result of the virus spread through networks of intravenous drug administration (IDU) (Poynard *et al.*, 2003; Pybus *et al.*, 2005; Thomson and Finch, 2005). The relative frequencies of HCV genotypes in different populations are not uniform, reflecting changes depending on the population and the age and route of transmission. The distribution of genotypes and subtypes has also changed over time. In industrialized countries, the subtype 1a is usually found in young individuals sharing IDU as a risk factor. On the other hand, subtype 1b is more common in older patients with a blood transfusions history (Pybus *et al.*, 2005). In Spain, the HCV prevalence is similar to other industrialized countries, being genotype 1 the most common followed by genotype 3 (Bruguera and Forns, 2006; Esteban *et al.*, 2008), although the distribution of subtypes presents some differences with other European countries (Lopez-Labrador *et al.*, 1997), with a high influence of subtype 1b infections. The distribution in Balearic archipelago is similar to the rest of the Spain country (Cifuentes *et al.*, 2004).

In most cases the infection progresses into a chronic state. The time between the beginning of chronic infection and a final resolution is very wide and it can range from 15 to about 40 years. Another feature is that chronic infection, often including the "some" or "mild to moderate" grade of cirrhosis, occurs without symptoms. For that reasons, the onset of infection is difficult to establish (Buti and Esteban, 2003). The ratio of spontaneous virus clearance (SVC) is around 15-30% of infected adults; in contrast, about

75-85% of patients achieve chronic disease. A higher rate of clearance is found in children and young women, reaching 45% of the total infected in this age range (Seeff, 2009).

Poynard *et al.* (2003) reviewed all known factors that may be associated with the development of fibrosis and, therefore, with the outcome of the disease. Age seems to be a key factor in the likelihood of spontaneous viral clearance because younger patients have higher spontaneous viral clearance rates. One explanation of this fact could be that the age is an indirect indicator of the strength of the immune response (Alter, 2006).

Men have worse prognosis than women in this infection, as well as being a habitual consumer of alcohol. Another group of variables that may be involved in the progress of the disease are related with the virus, such as, its genotype or viral load, although the evidence is still scarce. Patients with lower transaminases levels (ALT and AST) have lower fibrosis than those patients with elevated enzymes (Seeff, 2009). Another liver enzyme, the GGT, has been found in lower amounts in patients with better response rates, and in high levels in patients with hepatic steatosis and advanced liver fibrosis (Kau *et al.*, 2008). Levels of triglycerides and cholesterol in patients with chronic hepatitis C are lower than in patients who have spontaneously eliminated the virus. Moreover, patients with higher triglyceride levels, show a better response to treatment (Piver *et al.*, 2010).

For many years, the most effective treatment was based in the administration of interferon, producing a sustained viral response (SVR) of 20%. The combined use of interferon with ribavirin (RBV) and subsequent use of polyethylene glycolated interferon (pegIFN), represented a breakthrough in the treatment of disease caused by HCV, increasing the response ratio to 50%. Patients who do not decrease the viral load at first have worse prognosis to achieve an SVR than patients who quickly decrease viral load after treatment (Asselah *et al.*, 2010; Thomson and Finch, 2005). Patients infected with HCV genotype 2 or 3 achieve SVR in 80% of cases, while only 50% of patients infected with genotype 1 virus achieve SVR. Characteristics of the host such as older age, male gender, cirrhosis or hepatic steatosis presence, insulin resistance, the existence of diabetes or overweight are all related to a lower SVR. Coinfection with human immunodeficiency virus (HIV) or hepatitis B virus (HBV) are also associated with lower SVR (Poynard *et al.*, 2003; Heathcote, 2007; Asselah *et al.*, 2010; Rauch *et al.*, 2010).

An important aspect in clinical practice with patients infected with HCV is the monitoring of the treatment effectiveness based on repeated measurements of HCV RNA levels, in order to determine the efficacy and duration of the treatment. These

measurements of viral RNA levels should be performed before starting treatment, at 4, 12 and 24 weeks on treatment and 24 weeks after the end of the treatment to check whether patients have achieved SVR. The probability of achieving SVR is directly proportional to the time until the viral load becomes undetectable. The treatment should be discontinued at week 12 if the decrease is less than 2 logs (IU / mL) of viral load and at week 24 if there is still detectable viral load, as the probability that these patients get SVR is really low (between 1 and 3%; EASL, 2011).

It would be important to know the mechanisms that lie behind not responding patients before starting a treatment. The aim would be to anticipate what opportunities a patient has in each case to response. Recent studies have based these differences on the host genetic variability, trying to establish which host characteristics may give a better or worse prognosis. Different single nucleotide polymorphisms (SNP) have been related with HCV outcome. Ge *et al* (2009) found a strong relation between an SNP (rs12979860) close to the interleukin 28b gene (IL28b) and the treatment response. A large number of SNPs have been also studied to have some relation with the outcome, for example SNPs related to the immune system, as the studied in the Cytotoxic T-Lymphocyte Antigen 4 gene (CTLA4) (Yee *et al.*, 2003; Schott *et al.*, 2007; Nischalke *et al.*, 2010) or related with the lipid metabolism, as the studied in the low density lipoprotein receptor gene (LDLr) (Molina *et al.*, 2007; Ye, 2007; Mas Marques *et al.*, 2009). Other studies have used markers that could be associated with a worse prognosis of the disease, for example, the SNPs which determine other liver diseases as the mutations in HFE gene, among others (Geier *et al.*, 2004; Bonkovsky *et al.*, 2006). Finally, some studies have been performed with using populational markers as mitochondrial DNA (mtDNA) haplogroups, which can also be an indicator of the mitochondrial function (Hendrickson *et al.*, 2008; García-Álvarez *et al.*, 2011; Micheloud *et al.*, 2011).

The purpose of the present study was to relate different host SNPs in IL28b, CTLA4, LDLr, and HFE genes and mtDNA haplogroups, with the outcome of HCV infection in a Mallorcan coinfecting HCV-HIV population treated with pegIFN-RBV.

C. Materials and methods

Patient and control samples

Sixty-three HCV-HIV coinfecting patients were selected from the Infectious Diseases Unit of Son Llàtzer Hospital, Palma, Illes Balears (Spain). The study was approved by the local institutional ethics committee and all patients gave their informed consent to use their biological samples for medical research. Inclusion criteria were positive viraemia for both viruses (HCV and HIV), while seronegative for hepatitis B virus (HBV). Patients were treated with pegIFN-RIB except in SVC cases.

Three swabs were used to extract patient's oral cells and then left to dry and store at room temperature.

59 anonymous individuals (29 men and 30 women) belonging to Majorcan population, were randomly selected from DNA samples available in the Genetics Laboratory of the University of the Balearic Islands as controls. The polymorphisms described in this paper have never been studied in a coinfecting patient population and partially described in control population.

Variables

We built a database with patient variables selected retrospectively by chart review. The variables included basic characteristics of the patients (age, sex, body mass index, alcohol consumption), biochemical markers (AST -aspartate transaminase-, ALT -alanine transaminase-, cholesterol), blood count (haemoglobin, leukocytes, platelets), features related to infection (HCV genotype, HCV viral load, transmission route, HAART -Highly Active Antiretroviral Therapy) and variables related to effects of the disease (CD4 cell count, liver damage measured by Fibroscan and Biopsy, presence of hepatomegaly, splenomegaly and/or hepatic steatosis). Finally, variables related to HCV IFN and ribavirin combined treatment response (response at 4, 12 and 24 weeks, sustained viral response (SVR) and spontaneous clarification). The response was measured by means of viral load. Patients had a SVR when viral load was undetectable 24 weeks after finishing treatment. Spontaneous clearance was determined when treatment naive patients have negative viral load but were HCV seropositive.

Genomic polymorphisms

Seven SNPs were evaluated: IL28b, CTLA4+49, CTLA4-318, LDLr, HFE C282Y, HFE H63D and HFE S65C plus the Mitochondrial Haplogroup. In the control samples we

only analyse the polymorphisms that were not studied in Majorcan population: IL28b, both CTLA4, LDLr and HFE S65C (Table 1). For genotyping, we used mainly the RFLPs and, when it was necessary, the sequencing (LDLr and non-H haplogroups).

i. DNA extraction and amplification

A standard phenol-chloroform method and subsequent precipitation with ethanol and NaCl was used to extract DNA from oral swabs. The primers used to amplify IL28b polymorphism (427 bp) were 5'-GCTTATCGCATACGGCTAGG-3' (forward) and 5'-AGGGACCGCTACGTAAGTCA-3' (reverse). For the LDLr amplification (195 bp) we used 5'-TGGCAGAGACAGATGGTCAG-3' (forward) and 5'-CACTGTCCGAAGCCTGTTCT-3' (reverse). The primers used to amplify the CTLA4 (680 bp) and HFE polymorphisms (390 bp for C282Y and 208 bp for H63D/S65C) were described by Nischalke *et al.* (2010) and Feder *et al.* (1996), respectively. PCRs were carried out using 40ng of DNA, 0.2pM of each primer, 0.2mM of each dNTPs and 0.75 units of Taq polymerase. The amplification consisted in denaturation at 94°C for 10 minutes, followed by 35 cycles of denaturation at 94°C for 1 minute, 1 minute at the annealing temperature (59-60°C) and 1 minute of extension at 72°C. Ten minutes at 72°C was added to the end of the cycles.

RFLP motif -7025 AluI was analysed to determine haplogroup H using the following primers 5'-CCGTAGGTGGCCTGACTGGC-3' (forward) and 5'-TGATGGCAAATACAGCTCCT-3' (reverse). The PCR used 60ng of DNA, 2pM of each primer, 0.4mM of each dNTPs and 1.25 units of Taq polymerase. This haplogroup (124 bp) was detected with denaturation at 94°C for 5 minutes, followed by 35 cycles of denaturation at 94°C for 50 seconds, 1 minute 50 seconds at 59°C (annealing) and 1 minute of extension at 72°C. Five minutes at 72°C was added to the end of the cycles.

The mtDNA hypervariable regions of non-H samples were firstly amplified with denaturation at 96°C for 5 minutes, followed by 30 cycles of denaturation at 94°C for 1 minute, 1 minute at 50°C (annealing) and 1 minute of extension at 72°C. A 453-nt sequence in region I and 401-nt in region II of the mtDNA D-loop were analysed using the primers L15996 (5'-CTCCACCAATTAGCACCCAAAGC-3') and H16401 (5'-TGATTTACGGAGGATGGTG-3') in hypervariable region I and L48 (5'-CTCACGGGAGCTCTCCATGC-3') and H408 (5'-CTGTTAAAAGTGCATACCGCCA-3') in hypervariable region II. 80ng of DNA 0.2pM of each primer, 0.4mM of each dNTPs and 1 unit of Taq polymerase were used in each reaction.

ii. Genotyping by RFLPs

In Table 2a, we can see the restriction enzymes, the target point, the temperature and duration used in digestions. To design those not described in literature, we used the tools of <http://tools.neb.com/NEBcutter2/> website. Electrophoresis were carried out using 1X TAE buffer with medium agarose, and using ethidium bromide for staining DNA. Table 2a specifies the percentage of agarose used and Table 2b summarizes the band pattern results. To verify the new RFLPs, direct sequencing was performed, using the methodology described in 2.3.1.

iii. Genotyping by sequencing

The polymorphism located in LDLr gene and the non-H mtDNA haplogroups were detected by sequencing. The PCR products were purified using PCRapace Spin kit (Invitek, Berlin, Germany) and sequenced with BigDye [®] Terminator v.3.1. Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Finally, samples were precipitated with ethanol and analysed in an ABI 3130 automated sequencer (Applied Biosystems, Foster City, CA). The fragments were analysed using the Bioedit Sequence Alignment Editor software (Hall, 1999). Non-H mtDNA sequences were aligned and the most probable haplogroup was assigned by means of the Haplogrep software (haplogrep.uibk.ac.at/).

Statistical methodology

The genotypic and allelic frequencies of the polymorphisms were calculated by direct counting. Chi-square or Fisher tests, as appropriate, were performed on contingency tables to check differences in genotype distribution between patient and control groups.

To analyse associations with qualitative variables (sex, genotype virus, route of transmission, biopsy data, viral load, hepatomegaly, splenomegaly, and hepatic steatosis), Chi-square statistic or Fisher test were also used on contingency tables. Following the same procedure, the association between the genotypes and the response parameters (response at 4, 12 and 24 weeks, SVR and spontaneous clarification) was tested.

Quantitative variables were tested for a normal distribution by the Kolmogorov-Smirnov test. The variables with normal distribution were analysed using the *Student t* or the ANOVA, as comparing two or three groups, respectively. Quantitative variables that did not follow a normal distribution were analysed using the nonparametric tests, Mann-Whitney or Kruskal-Wallis, as comparing two or three groups, respectively. The statistical

calculations were performed using SPSS for Windows (Rel. 15.0.1 2006 Chicago: SPSS Inc.).

D. Results

Baseline characteristics of the subjects

The main features of the studied cohort are shown in Table 3. Fibrosis stage was obtained by biopsy (grouped in F0-F1 stages) and Fibroscan (data below 7.2Kpa were also considered as F0-F1 stage). The variables with normal distribution were body mass index, cholesterol, platelets, haemoglobin, leukocytes, Fibroscan and age, while quantitative variables that did not follow a normal distribution were AST, ALT, HCV viral load and the absolute CD4 count.

Sex was not distributed evenly as male patients more than doubled female patients. Males presented higher levels of AST, ALT than females and present IDU as the main route of transmission, but this fact is not reflected in fibrosis levels. Alcoholic consumption is present only in a few female patients. No differences between males and females were found in the other variables.

Frequencies of the polymorphisms

The genetic polymorphism of the patients were unequivocally assigned in all cases (n=63). Six out of 63 patients showed spontaneous clearance, so they were not treated with pegIFN-RIB therapy. IL28b, LDLr, CTLA4 and S65C polymorphisms were analysed in 53-58 individuals in the control group, depending on the SNP. The distribution of polymorphisms is shown in Table 4. All of them are in Hardy-Weinberg equilibrium and the SNPs variation is not correlated. No significant differences were found between males and females in any case (data not shown). Differences between patients and controls were only detected in H63D mutation, being higher in the group of patients. Mitochondrial haplogroups were grouped in four principal groups H, J, U and a group with the remaining of haplogroups.

The analyses concerning IL28b polymorphism showed that patients with CC genotype have lower levels of leukocytes ($p=0.047$) and platelets ($p=0.004$). Although in the former, an outlier patient with high levels makes that this result occurs and the effect is diluted when this patient is removed. On the other hand, patients with this genotype are those that are less infected by genotype 4 of the virus ($p=0.049$). All these effects are most

evident when we grouped heterozygous TC and homozygous TT patients and we compare them with CC genotype.

Patients carriers of G allele in CTLA4+49 polymorphism present ALT and AST transaminase levels substantially higher ($p=0.021$ and $p=0.063$, respectively). Patients carrying the CC genotype of CTLA4-318 present Fibroscan levels above average (data not shown), although it is not reflected in the pooled fibrosis. Patients with the AA genotype in LDLr have higher proportion of hepatic steatosis ($p=0.047$) and a lower body mass index ($p=0.029$). We found no homozygous patient for the C282Y mutation and the S65C prevalence of the mutation is extremely low in both controls and patients groups. Besides the difference in prevalence of H63D mutation between patients and controls, the most remarkable feature is that 3 of the four C282Y/wt, the H63D/H63D and the S65C/wt patients were subjected to biopsy, so it could be an evidence of liver involvement. Hepatomegaly is present in patients belonging to most major haplogroups H, J and U, while the ALT transaminase was significantly higher in patients with haplogroup H and higher, but not significant, in carriers of haplogroup J.

Response to pegIFN-RIB treatment

Data were obtained for treatment response at 4, 12 and 24 weeks, and finally for sustained viral response (SVR). Data were available in 55 patients. 31 (56.4%) of them have not achieved SVR. Six (9.5%) of the studied patients showed spontaneous clearance (Table 5). Patients infected with genotype 3 virus showed significantly higher SVR than patients infected with genotype 1 and 4 ($p= 0.008$).

The results concerning the relation between the treatment response and the polymorphism analyses are shown in Table 6. The most significant results were obtained in the IL28b polymorphism. There is a better response and therefore a negative viral load in patients with CC genotype. Difference becomes more evident when they are compared with the pool of CT and TT patients (Figure 1a). The spontaneous clearance is higher in patients carrying the G allele in the CTLA4+49 polymorphism ($p= 0.02$) (Figure 1b). On the other hand, without being statistically significant, patient with T allele in CTLA4-318 SNP show higher SVR.

Patients with no SVR ($n=31$; 56.4%) were divided in two groups. In the first group, we included patient who never completely eliminated the virus ($n=17$; 30.9%) and other group was formed by patients with negative viral load at some point after starting the treatment but with a later spread of the infection ($n=14$; 25.5%). The latter were

considered as relapsers. If we analyze the relationship between IL28b polymorphism genotype and relapsers, we found that only one patient with CC genotype had never eliminated the virus. Otherwise, over 50% of patients who had responded to treatment belong to this genotype ($p=0.05$) (Figure 1a).

Analysing the relation between other variables and the response to treatment, AST and ALT levels were higher in patients with SVR ($p=0.001$ and $p=0.002$, respectively) (data not shown).

E. Discussion

The response to the viral treatment of the patients depends on viral and host factors. In relation to host factors, the role of polymorphisms that could affect the response has been studied. In this way, we have investigated the results obtained in the analyses of different polymorphisms with the response to pegIFN-rib treatment in patients coinfecting with HCV and HIV in Majorcan population.

The strong relation found between an SNP (rs12979860) close to the interleukin 28b gene (IL28b) and the treatment response (Ge *et al.*, 2009) indicates the potential importance of this polymorphism in the evolution of the illness. Although the frequency detected in our cohort is similar to the European American cohort analysed by Ge *et al.* (2009), the C allele frequency is lower than in other Spanish cohorts (Montes-Cano *et al.*, 2010; Rallón *et al.*, 2010). Patients and controls showed similar frequencies in IL28b polymorphism. The frequencies of other SNPs related to the immune system as the CTLA4 gene (Yee *et al.*, 2003; Schott *et al.*, 2007; Nischalke *et al.*, 2010) or to the lipid metabolism, as the LDLr gene (Molina *et al.*, 2007; Ye, 2007; Mas Marques *et al.*, 2009) are similar to those found in other populations (Yee *et al.*, 2003; Mas Marques *et al.*, 2009) and, as before, no differences were found in the frequencies between patient and control cohorts. In the same way, no differences were found in the mitochondrial haplogroups. On the contrary, H63D mutation in HFE gene has shown differences between the cohort of patients and the control population. It is noteworthy that Spain is the country with highest frequency of the H63D mutation (Guix *et al.*, 2002).

It is known that the analyses between SVR and the HCV genotypes indicate that patients infected with genotypes 1 and 4 are related to a worst treatment response, whereas those infected by genotype 3 respond better (Rauch *et al.*, 2010). The average response of

our patient cohort is similar to that found in the literature, although the ratio of SVC was slightly lower (Seeff, 2009).

What relationship can we find, in the present study, between the genetic polymorphisms of the host and the response to treatment? IL28b polymorphism has been the most helpful, along with the traditional use of viral genotypes to anticipate the response of a patient with treatment (Rauch *et al.*, 2010). We have detected that the frequency of CC genotype is very low in patients with genotype 4. Carriers of CC genotype respond better, but patients infected with genotype 4 are among those with worst answer, for this reason it is difficult to establish which of the two effects is more important. These interactions have to be taken into account to determine whether or not a polymorphism improves the response to the treatment. The most significant data are found in response at 4 weeks, so it seems that the CC genotype promote a rapid response to treatment. All results are more evident when we compare homozygotes CC with a group formed by the heterozygotes and the homozygotes TT. From the whole cohort, only one IL28b CC genotype patient infected by genotype 3 virus has not achieved SVR. On the other hand, only one IL28b TT genotype patient infected by genotype 1 or 4 virus has achieved SVR. The analyses of patients who have relapsed shows that only one patient who had never responded to treatment presents the CC genotype. This indicates the importance of this polymorphism in a first response to infection rather than a sustained viral response.

On the other hand, no relationship between the LDLr gene and the response to treatment was found. Several publications showed that the GG genotype showed a better response (Hennig *et al.*, 2002; Pineda *et al.*, 2011), but in our cohort only a relation between AA genotype and an increased hepatic steatosis was found.

It is noteworthy that the SVC was more evident in patients with the G allele in CTLA4+49 polymorphism, because the 6 individuals in the cohort who presented spontaneous clearance, have this allele. This result is in agreement with the findings in Yee *et al.* (2003), where this allele improves the treatment response in genotype 1 HCV monoinfected patients treated with non-pegylated interferon, although our patients do not get the same results in relation with the treatment response. Perhaps, this is due to the different characteristics of our cohort, coinfecting with HIV and with HCV genotypes 1, 3 and 4 represented. Other studies concluded that GG genotype was associated with better response, but not with the spontaneous clarification (Nischalke *et al.*, 2010). It does not

seem that patients with the G allele have a trend to respond better to treatment, in any case.

Mutations in the HFE gene seem to have no connection with any variable studied but it has to be noted that all H63D homozygotes, all S65C heterozygotes and almost all C282Y heterozygotes have been subjected to biopsy. This could indicate that these patients had some evidence that forced him to perform these tests, although it is not reflected in the pooled fibrosis.

In conclusion, this study wanted to highlight the importance of knowing the genetic background of the patients before starting a treatment taking into account the specific characteristics of the cohort. The results showed the relevance of CC genotype in IL28B polymorphism in the predisposition to a better response to pegIFN-RIB treatment, and also, a greater susceptibility to spontaneous viral clearance in G allele carriers in CTLA4+49 polymorphism.

F. Acknowledgements

This work was supported by grants 4326/2007 and 6557/2010 from the Direcció General de Recerca, Desenvolupament Tecnològic i Innovació. Conselleria d'Economia, Hisenda i Innovació, Govern Balear, Spain; projects PI10/00512, PI10/01734, and CIBEResp, Instituto de Salud Carlos III, Spain; projects BFU2008 03000BMC and BFU2011-24112 from Ministerio de Ciencia e Innovación, Spain. F.X.L. holds a P.I. position supported by the Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Spanish Ministry of Science.

G. References

-
- Alter, H., 2006. Viral hepatitis. *Hepatology* 43, S230-234.
- Alter, M.J., 2007. Epidemiology of hepatitis C virus infection. *World J. Gastroenterol.* 13, 2436-2441.
- Asselah, T., Estrabaud, E., Bieche, I., Lapalus, M., De Muynck, S., Vidaud, M., Saadoun, D., Soumelis, V., Marcellin, P., 2010. Hepatitis C: viral and host factors associated with non-response to pegylated interferon plus ribavirin. *Liver Int.* 30, 1259-1269.

- Bonkovsky, H.L., Naishadham, D., Lambrecht, R.W., Chung, R.T., Hoefs, J.C., Nash, S.R., Rogers, T.E., Banner, B.F., Sterling, R.K., Donovan, J.A., Fontana, R.J., Di Bisceglie, A.M., Ghany, M.G., Morishima, C., Group, H.-C.T., 2006. Roles of iron and HFE mutations on severity and response to therapy during retreatment of advanced chronic hepatitis C. *Gastroenterology* 131, 1440-1451.
- Bruguera, M., Forns, X., 2006. Hepatitis C en España. *Med. Clin. (Barc)*. 127, 113-117.
- Buti, M., Esteban, R., 2003. VII International Symposium on Viral hepatitis. *Acción Médica*, Madrid.
- Cifuentes, C., Juan, M., Sastre, M., Payeras, A., Fernández-Baca, V., Homar, F., Gallegos, C., Bassa, A., 2004. Coinfected by HIV-HCV: epidemiological data. First genotype determination in a sample from Mallorca Island., 27th World Congress of Internal Medicine, Granada, Spain.
- EASL (2011). EASL Clinical Practice Guidelines: management of hepatitis C virus infection. *J. Hepatol.* 55, 245-264.
- Esteban, J.I., Sauleda, S., Quer, J., 2008. The changing epidemiology of hepatitis C virus infection in Europe. *J. Hepatol.* 48, 148-162.
- Feder, J.N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D.A., Basava, A., Dormishian, F., Domingo, R., Jr., Ellis, M.C., Fullan, A., Hinton, L.M., Jones, N.L., Kimmel, B.E., Kronmal, G.S., Lauer, P., Lee, V.K., Loeb, D.B., Mapa, F.A., McClelland, E., Meyer, N.C., Mintier, G.A., Moeller, N., Moore, T., Morikang, E., Prass, C.E., Quintana, L., Starnes, S.M., Schatzman, R.C., Brunke, K.J., Drayna, D.T., Risch, N.J., Bacon, B.R., Wolff, R.K., 1996. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat. Genet.* 13, 399-408.
- García-Álvarez, M., Guzmán-Fulgencio, M., Berenguer, J., Micheloud, D., Campos, Y., López, J.C., Cosín, J., Miralles, P., Alvarez, E., Resino, S., 2011. European mitochondrial DNA haplogroups and liver fibrosis in HIV and hepatitis C virus coinfecting patients. *AIDS* 25, 1619-1926.
- Ge, D., Fellay, J., Thompson, A.J., Simon, J.S., Shianna, K.V., Urban, T.J., Heinzen, E.L., Qiu, P., Bertelsen, A.H., Muir, A.J., Sulkowski, M., McHutchison, J.G., Goldstein, D.B., 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461, 399-401.
- Geier, A., Reugels, M., Weiskirchen, R., Wasmuth, H.E., Dietrich, C.G., Siewert, E., Gartung, C., Lorenzen, J., Bosserhoff, A.K., Brugmann, M., Gressner, A.M., Matern, S., Lammert, F., 2004. Common heterozygous hemochromatosis gene mutations are risk factors for inflammation and fibrosis in chronic hepatitis C. *Liver Int.* 24, 285-294.

- Guix, P., Picornell, A., Parera, M., Galmes, A., Obrador, A., Ramon, M.M., Castro, J.A., 2002. Distribution of HFE C282Y and H63D mutations in the Balearic Islands (NE Spain). *Clin. Genet.* 61, 43-48.
- Hall, T.A., 1999. Bioedit: a user friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95-98.
- Heathcote, E.J., 2007. Antiviral therapy: chronic hepatitis C. *J. Viral Hepatitis* 14, 82-88.
- Hendrickson, S.L., Hutcheson, H.B., Ruiz-Pesini, E., Poole, J.C., Lautenberger, J., Sezgin, E., Kingsley, L., Goedert, J.J., Vlahov, D., Donfield, S., Wallace, D.C., O'Brien, S.J., 2008. Mitochondrial DNA haplogroups influence AIDS progression. *AIDS* 22, 2429-2439.
- Hennig, B.J., Hellier, S., Frodsham, A.J., Zhang, L., Klenerman, P., Knapp, S., Wright, M., Thomas, H.C., Thursz, M., Hill, A.V., 2002. Association of low-density lipoprotein receptor polymorphisms and outcome of hepatitis C infection. *Genes Immun.* 3, 359-367.
- Kau, A., Vermehren, J., Sarrazin, C., 2008. Treatment predictors of a sustained virologic response in hepatitis B and C. *J. Hepatol.* 49, 634-651.
- López-Labrador, F.X., Ampurdanés, S., Forn, X., Castells, A., Sáiz, J.C., Costa, J., Bruix, J., Sánchez Tapias, J.M., Jiménez de Anta, M.T., Rodés, J., 1997. Hepatitis C virus (HCV) genotypes in Spanish patients with HCV infection: relationship between HCV genotype 1b, cirrhosis and hepatocellular carcinoma. *J. Hepatol.* 27, 959-965.
- Lynas, C., 1997. A cheaper and more rapid polymerase chain reaction-restriction fragment length polymorphism method for the detection of the HLA-H gene mutations occurring in hereditary hemochromatosis. *Blood* 90, 4235-4236.
- Mas Marques, A., Mueller, T., Welke, J., Taube, S., Sarrazin, C., Wiese, M., Halangk, J., Witt, H., Ahlenstiel, G., Spengler, U., Goebel, U., Schott, E., Weich, V., Schlosser, B., Wasmuth, H.E., Lammert, F., Berg, T., Schreier, E., 2009. Low-density lipoprotein receptor variants are associated with spontaneous and treatment-induced recovery from hepatitis C virus infection. *Infect. Genet. Evol.* 9, 847-852.
- Micheloud, D., Berenguer, J., Guzmán-Fulgencio, M., Campos, Y., García-Álvarez, M., Catalán, P., Cosín, J., Miralles, P., López, J.C., Resino, S., 2011. European mitochondrial DNA haplogroups and metabolic disorders in HIV/HCV-coinfected patients on highly active antiretroviral therapy. *J. Acquir. Immune Defic. Syndr.* 58, 371-378.
- Molina, S., Castet, V., Fournier-Wirth, C., Pichard-García, L., Avner, R., Harats, D., Roitelman, J., Barbaras, R., Graber, P., Ghersa, P., Smolarsky, M., Funaro, A.,

- Malavasi, F., Larrey, D., Coste, J., Fabre, J.M., Sa-Cunha, A., Maurel, P., 2007. The low-density lipoprotein receptor plays a role in the infection of primary human hepatocytes by hepatitis C virus. *J. Hepatol.* 46, 411-419.
- Montes-Cano, M.A., García-Lozano, J.R., Abad-Molina, C., Romero-Gomez, M., Barroso, N., Aguilar-Reina, J., Nunez-Roldan, A., González-Escribano, M.F., 2010. Interleukin-28B genetic variants and hepatitis virus infection by different viral genotypes. *Hepatology* 52, 33-37.
- Nischalke, H.D., Vogel, M., Mauss, S., Baumgarten, A., Lutz, T., Danta, M., Naumann, U., Coenen, M., Sauerbruch, T., Rockstroh, J.K., Spengler, U., Nattermann, J., 2010. The cytotoxic lymphocyte antigen 4 polymorphisms affect response to hepatitis C virus-specific therapy in HIV(+) patients with acute and chronic hepatitis C virus co-infection. *AIDS* 24, 2001-2007.
- Picornell, A., Gómez-Barbeito, L., Tomàs, C., Castro, J.A., Ramon, M.M., 2005. Mitochondrial DNA HVRI variation in Balearic populations. *Am. J. Phys. Anthropol.* 128, 119-130.
- Pineda, J.A., Caruz, A., Di Lello, F.A., Camacho, A., Mesa, P., Neukam, K., Rivero-juárez, A., Macías, J., Gomez-Mateos, J., Rivero, A., 2011. Low-density lipoprotein receptor genotyping enhances the predictive value of IL28B genotype in HIV/hepatitis C virus-coinfected patients. *AIDS* 25, 1415-1420.
- Piver, E., Roingeard, P., Pagès, J.C., 2010. The cell biology of hepatitis C virus (HCV) lipid addiction: molecular mechanisms and its potential importance in the clinic. *Int. J. Biochem. Cell Biol.* 42, 869-879.
- Poynard, T., Yuen, M.F., Ratziu, V., Lai, C.L., 2003. Viral hepatitis C. *Lancet* 362, 2095-2100.
- Pybus, O.G., Cochrane, A., Holmes, E.C., Simmonds, P., 2005. The hepatitis C virus epidemic among injecting drug users. *Infect. Genet. Evol.* 5, 131-139.
- Rallón, N.I., Naggie, S., Benito, J.M., Medrano, J., Restrepo, C., Goldstein, D., Shianna, K.V., Vispo, E., Thompson, A., McHutchison, J., Soriano, V., 2010. Association of a single nucleotide polymorphism near the interleukin-28B gene with response to hepatitis C therapy in HIV/hepatitis C virus-coinfected patients. *AIDS* 24, F23-29.
- Rauch, A., Kutalik, Z., Descombes, P., Cai, T., Di Iulio, J., Mueller, T., Bochud, M., Battegay, M., Bernasconi, E., Borovicka, J., Colombo, S., Cerny, A., Dufour, J.F., Furrer, H., Gunthard, H.F., Heim, M., Hirschel, B., Malinverni, R., Moradpour, D., Mullhaupt, B., Witteck, A., Beckmann, J.S., Berg, T., Bergmann, S., Negro, F., Telenti, A., Bochud, P.Y., Swiss Hepatitis, C.C.S., Swiss, H.I.V.C.S., 2010. Genetic variation in IL28B is associated with chronic hepatitis C and treatment

- failure: a genome-wide association study. *Gastroenterology* 138, 1338-1345, 1345 e1331-1337.
- Schott, E., Witt, H., Hinrichsen, H., Neumann, K., Weich, V., Bergk, A., Halangk, J., Muller, T., Tinjala, S., Puhl, G., Neuhaus, P., Wiedenmann, B., Berg, T., 2007. Gender-dependent association of CTLA4 polymorphisms with resolution of hepatitis C virus infection. *J. Hepatol.* 46, 372-380.
- Seeff, L.B., 2009. The history of the "natural history" of hepatitis C (1968-2009). *Liver Int.* 29, 89-99.
- Thomson, B.J., Finch, R.G., 2005. Hepatitis C virus infection. *Clin. Microbiol. Infect.* 11, 86-94.
- Ye, J., 2007. Reliance of host cholesterol metabolic pathways for the life cycle of hepatitis C virus. *PLoS Pathog.* 3, e108.
- Yee, L.J., Perez, K.A., Tang, J., van Leeuwen, D.J., Kaslow, R.A., 2003. Association of CTLA4 polymorphisms with sustained response to interferon and ribavirin therapy for chronic hepatitis C virus infection. *J. Infect. Dis.* 187, 1264-1271.

H. Figures and Tables

Table 1. Host genetic polymorphisms analysed

Gene	Polymorphism	Nucleotide change
Interleukin 28b (IL28b)	rs12979860	T / A
Cytotoxic T-Lymphocyte Antigen 4 (CTLA4)	+49 (rs231775)	G / A
	-318 (rs5742909)	C / T
Hereditary Hemochromatosis gene (HFE)	C282Y	G / A
	H63D	C / G
	S65C	T / A
Low Density Lipoprotein receptor (LDLr)	rs14158	A / G
Mitochondrial DNA haplogroup	H	
	Others	

Table 2. RFLP analyses: (a) conditions, (b) band pattern obtained

(a)

Polymorphisms	Enzymes	Target point (polymorphic nucleotide)	Temperature and time	% agarose
IL28b rs12979860	<i>Hpy</i> 166II	GTN [^] NAC	37°C – 3h	2%
CTLA4 +49	<i>Fnu</i> 4HI	GC [^] NGC	37°C – 2h	4%
CTLA4 -318	<i>Mn</i> I	CCTC(N) ₇ [^]	37°C – 2h	4%
HFE C282Y	<i>Rsa</i> I (Lynas, 1997)	GT [^] AC	37°C – 3,5h	3%
HFE H63D	<i>Bcl</i> I (Lynas, 1997)	T [^] GATCA	50°C – 2h	3%
HFE S65C	<i>Hin</i> fI	G [^] ANTC	37°C – 3h	3%
Haplogrup H	<i>Alu</i> I	AG [^] CT	37°C – 3h	2%

(b)

	Motiv	bp fragment	Motiv	bp fragment	Unspecific fragment (bp)
IL28b rs12979860	C	320	T	28, 292	
CTLA4 +49	A	223	G	28, 195	457
CTLA4 -318	C	59, 71	T	130	53, 111, 376
HFE C282Y	Normal	250	Mutated	110, 140	
HFE H63D	Normal	70, 138	Mutated	208	
HFE S65C	Normal	58, 150	Mutated	208	
Haplogrup H	Haplo H	124	Non H	80	

Table 3. Baseline demographic, epidemiological and clinical characteristics of HIV–HCV-co-infected subjects. Number of affected or mean and standard deviation (SD) are indicated in the variables.

			Male (n=43)	Female (n=20)	Total (n=63)	p	
<i>Mean (SD)</i>	AGE, years	na = 0	42 (5)	41 (6)	42 (5)		
	BMI	na = 4	23.84 (2.66)	24.23 (3.86)	23.97 (3.09)		
	Cholesterol, mg/dL	na = 2	173 (38)	189 (44)	178 (40)		
	AST level, U/mL	na = 1	68 (43)	51 (45)	63 (44)	p<0.05	
	ALT level, U/mL	na = 1	87 (62)	55 (38)	77 (57)	p<0.05	
	CD4 cell count, cell/mm ³	na = 1	502 (299)	543 (280)	515 (291)		
	Haemoglobin levels, g/dL	na = 11	14.90 (1.74)	13.59 (0.83)	14.44 (1.61)	p<0.005	
	Platelet count, 10 ⁹ /L	na = 1	193.28 (52.55)	217.79 (73.139)	200.79 (60.08)		
	Leukocytes, 10 ³ /mm ³	na = 12	6.16 (1.75)	5.12 (1.57)	5.81 (1.75)		
<i>N (%)</i>	HCV viral load	Negative (SVC)	na = 0	2	4	6 (9.5%)	
		Low		27	10	37 (58.7%)	
		High		14	6	26 (31.7%)	
	HCV genotypes	1		18	6	24 (42.9%)	
		3	na = 7	15	4	19 (33.9%)	
		4		8	5	13 (23.2%)	
	Fibrosis stage (Metavir)	F0-F1	na = 10	10	6	16 (30.2%)	
		F2-F4		27	10	37 (69.8%)	
	Alcohol consumption	Yes		0	3	3 (5.1%)	
		No	na = 4	35	13	48 (81.4%)	p<0.05
		Abstinent		6	2	8 (13.6%)	
	Route of Transmission	Intravenous drug use	na = 3	37	13	50 (83.3%)	p<0.05
Other			4	6	10 (16.7%)		
Highly active antiretroviral therapy	Yes		39	16	55 (90.2%)		
	No	na = 0	3	3	6 (9.8%)		
Hepatomegaly	Yes	na = 6	9	2	11 (19.3%)		
	No		32	14	46 (80.7)		
Hepatic steatosis	Yes	na = 18	9	1	10 (22.2)		
	No		22	13	35 (77.8)		
Splénomegaly	Yes		10	4	14 (24.6)		
	No	na = 6	31	12	43 (75.4)		

BMI: body mass index. AST: aspartate transaminase. ALT: alanine transaminase. Alcohol consumption: Yes: >50g for >2 years; Abstinent: no alcohol consumption for 6 month. HCV viral load: High: >600000 IU/mL. SVC: Spontaneous Vrial Clearance. Routes of transmissions: heterosexual or homosexual contact and transfusions. Fibrosis stage: data from biopsy and Fibroscan (F0-F1 and < 7.2Kpa).

na: not available data

Table 4. Polymorphism distribution in patient and control cohorts.

Polymorphisms		Patients n (%)	Controls n (%)
IL28b rs12979860	CC	23 (36.5)	19 (35.8)
	CT	31 (49.2)	27 (50.9)
	TT	9 (14.3)	7 (13.2)
CTLA4 +49	AA	29 (46.0)	30 (54.6)
	AG	28 (44.4)	18 (32.7)
	GG	6 (9.6)	7 (12.7)
CTLA4 -318	CC	55 (87.3)	51 (87.9)
	CT	7 (11.1)	6 (10.4)
	TT	1 (1.6)	1 (1.7)
HFE C282Y	N	59 (93.7)	181 (94.3) ¹
	Hz	4 (6.3)	11 (5.7) ¹
HFE H63D	N	25 (39.7)	120 (62.5) ¹
	Hz	34 (54.0)	69 (35.9) ¹
	Ho	4 (6.3)	3 (1.6) ¹
HFE S65C	N	60 (96.2)	51 (96.2)
	Hz	3 (4.8)	2 (3.8)
LDLr	GG	33 (52.4)	28 (51.8)
	AG	24 (38.1)	19 (35.2)
	AA	6 (9.5)	7 (13)
mtDNA Haplogroups	H	28 (44.4)	14 (32.6) ²
	J	10 (15.9)	4 (9.3) ²
	U	8 (12.7)	6 (14.0) ²
	other	17 (27.0)	19 (44.2) ²

¹: Data from Guix *et al.* (2002). ²: Data from Picornell *et al.* (2005)

Ho = homozigotes; Hz = heterozigotes

Table 5. Distribution of the Response to pegIFN-RIB treatment in males and females of the patient cohort.

N (%)		Male	Female	Total
VR4	Negative	14	4	18 (40.9%)
	Positive	19	7	26 (59.1%)
VR12	Negative	30	8	38 (69.1%)
	< 2 log	3	1	4 (7.3)
	Positive	8	5	13 (23.6)
VR24	Negative	31	8	39 (83.0%)
	Positive	5	3	8 (17.0%)
SVR	No	19	12	31 (56.4%)
	Yes	20	4	24 (43.6%)
SVC	No	41	16	57 (90.5%)
	Yes	2	4	6 (9.5%)

VR4, VR12, VR24: Viral load after 4, 12 and 24 weeks.

Negative: non detectable viral load; < 2 log: viral load has decreased more than two logarithms.

Positive: viral load has not significantly changed.

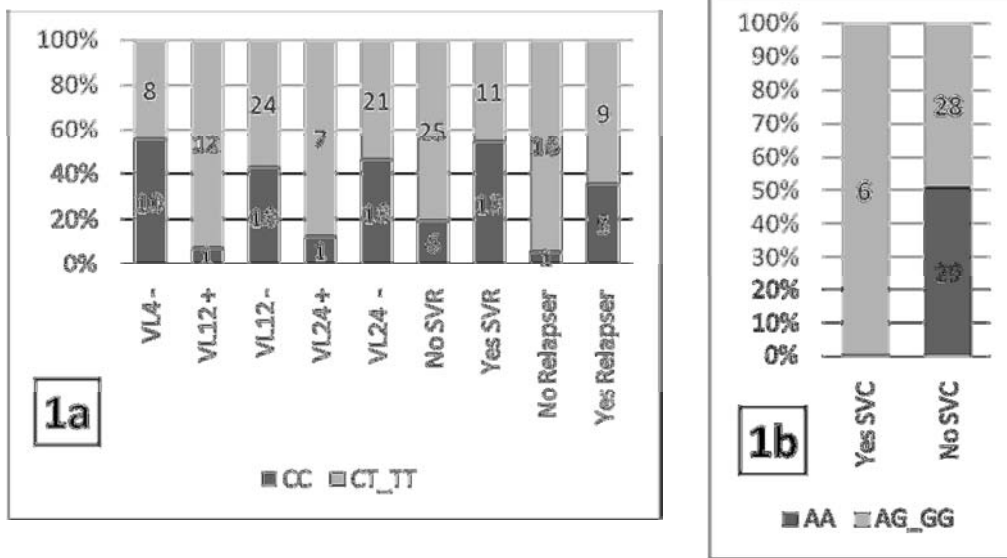
SVR: Sustained viral response after 24 weeks after treatment.

SVC: Spontaneous viral clearance.

Table 6. Statistical analyses (Chi-square or Fisher tests) of the response to pegIFN-RIB treatment when we analysed the differences between the genotypes ("/" means that the genotypes have been compared, and "-" that the genotypes are grouped). Values indicated the significant probabilities. Lack of signification is indicated as "-". The analyses were done between patients who have responded, or not, at 4 (VR4), 12 (VR12), 24 (VR24) weeks after starting the treatment and patients who have achieved, or not, spontaneous viral clearance (SVC) or sustained viral response (SVR).

Polymorphisms		VR4	VR12	VR24	SVR	SVC
IL28b rs12979860	CC / CT / TT	0.015	0.057	0.098	0.024	-
	CC-CT / TT	0.076	-	-	-	-
	CC / CT-TT	0.006	0.036	0.081	0.008	-
CTLA4 +49	AA / AG / GG	-	-	-	-	0.040
	AA-AG / GG	-	-	-	-	-
	AA / AG-GG	-	-	-	-	0.020
CTLA4 -318	CC / CT / TT	-	-	-	0.074	-
	CC-CT / TT	-	-	-	-	-
	CC / CT-TT	-	-	-	0.061	-
HFE C282Y	N / Hz	-	-	-	-	-
HFE H63D	N / Hz / Ho	-	-	-	-	-
HFE S65C	N / Hz	-	-	-	-	-
LDLr	AA / AG / GG	-	-	-	-	-
	AA-AG / GG	-	-	-	-	-
	AA / AG-GG	-	-	-	-	-
mtDNA Haplogroups	H / J / U / other	-	-	-	-	-

Figure 1. (1a) Grouped IL28b genotype (CC and CT_TT) and response to treatment. (1b) Grouped CTLA4+49 genotype (AA and AG_GG) and spontaneous viral clearance. The number of patients is indicated in the graphics.



(-) Negative: non detectable viral load or viral load has decreased more than two logarithms.
 (+) Positive: viral load has not significantly changed.

L'objectiu principal d'aquesta tesi ha estat relacionar factors genètics, epidemiològics i clínics, tant del virus com de l'hoste amb el desenvolupament de la malaltia produïda per la coinfecció del virus de l'hepatitis C (HCV) i del virus de la immunodeficiència humana (VIH), en dues cohorts de pacients procedents de l'hospital Son Llàtzer (Mallorca).

La consecució d'aquest objectiu es va fer mitjançant dues aproximacions que es discuteixen a continuació:

1. ESTUDI DEL GENOMA VÍRIC *versus* DANY HEPÀTIC

En el primer estudi, del qual derivaren dues publicacions, es va dur a terme una estimació dels factors potencialment involucrats amb el dany hepàtic utilitzant la metodologia GLM d'una forma semblant a Szmargd *et al.* (2006), per tal d'identificar variables clíniques dels hostes i característiques genètiques del VHC que poguessin estar relacionades amb un pitjor comportament de la malaltia produïda pel virus en pacients coinfectats amb HCV i HIV. Aquesta metodologia va permetre la combinació d'eines filogeogràfiques amb el modelatge estadístic de variables qualitatives i quantitatives i combinar les anàlisis de les característiques demogràfiques, epidemiològiques i clíniques dels pacients conjuntament amb dades del genoma del virus en una única anàlisi. La metodologia GLM permet fer aquestes anàlisis d'una manera directa, ja que és capaç de detectar estadísticament aquelles variables que aporten més significació al model seleccionat (Venables i Ripley, 1999). Els models resultants permeten reduir, per tant, el nombre de variables predictives en les anàlisis finals a aquelles més significatives, alhora que s'elegí el millor model que descrivia la variable resposta, en el nostre cas, el dany hepàtic.

Una de les formes de detectar aquest dany hepàtic és l'avaluació del nivell de fibrosi mitjançant la biòpsia hepàtica. Això moltes vegades no és possible o no està disponible en estudis retrospectius i/o d'una cohort anònima, com el nostre cas, on la informació directa sobre fibrosi hepàtica és reduïda. Per defugir d'aquesta limitació, es van emprar tres índexs àmpliament utilitzats en la bibliografia: l'índex de Forns (IF), l'APRI i el FIB-4 (Wai *et al.*, 2003), que estimen el dany hepàtic produït per la malaltia de forma indirecta (Forns *et al.*, 2002; Sterling *et al.*, 2006; Vallet-Pichard *et al.*, 2007).

És conegut que aquests índexs poden donar resultats no del tot acurats, de fet, semblen més acurats per pacients coinfectats amb VHC i VIH que en pacients monoinfectats (Ramos Paesa *et al.*, 2007), malgrat el fet que els nivells de plaquetes o transaminases (emprats per calcular els índexs) poden veure's alterats pel mateix VIH o pel tractament antiretroviral utilitzat per combatre'l. Aquesta falta de precisió es va comprovar quan es compararen els tres índexs amb els resultats de les dades de biòpsies de les que es disposava. La correlació mitjana estava al voltant del 30%, donant suport a la reduïda fiabilitat dels índexs. No podem descartar l'influència que pot tenir sobre la baixa correlació el limitat nombre de biòpsies realitzades en aquesta cohort, només 9 de 37 pacients, que podien produir un biaix important. Per una altra banda, els tres índexs mostraren una alta correlació entre ells (al voltant del 86% de mitjana), i aquesta correlació es va veure reflectida en els anàlisis GLM, no només quan es feren per separat, si no també quan s'utilitzaren com una única variable resposta.

Tal com s'ha donat a entendre, els resultats han de ser interpretats amb cura. No obstant això, alguns autors han posat damunt la taula la qüestió de si les biòpsies hepàtiques suposen un gran valor per l'avaluació i diagnòstic de la fibrosi, degut a errors de mostratge (Bedossa *et al.*, 2003), encara que avui en dia, la detecció de fibrosi utilitzant la biòpsia segueix considerant-se *l'estàndard d'or*. Malgrat això, el fet de que sigui un procediment invasiu i amb possibles complicacions fa, alhora, que sigui desitjable trobar variables substituïdes per avaluar la fibrosi o emprar altres tècniques no invasives com és el cas de l'elastografia.

Per tal d'incrementar la precisió de les mesures indirectes de fibrosi, es van utilitzar els diferents índexs simultàniament en les anàlisis com a una única variable de resposta. A l'estudi, que es va realitzar utilitzant les seqüències de la regió *Core* del VHC, es varen utilitzar els tres índexs (IF, APRI i FIB-4) per a realitzar el modelatge estadístic, mentre que a l'estudi realitzat utilitzant les seqüències de la regió *NS5B*, la metodologia es va millorar quan s'utilitzaren només els índexs APRI i FIB-4, per tal de corregir el biaix que es podria produir considerant índexs molt relacionats. En les anàlisis varen triar la distribució Gaussiana, en lloc de multinomial (com s'indicà a Szmargd *et al.* (2006)), ja que els índexs de fibrosi presentaven un ventall molt ampli de valors.

Per determinar les característiques genètiques del VHC, es seqüenciaren dues regions del virus, la regió *Core* i la regió *NS5B* que s'utilitzaren en anàlisis independents. La

regió *Core* s'utilitzà en l'anàlisi usant directament la seva seqüència de nucleòtids, mentre que la seqüència de la regió *NS5B* es va traduir a aminoàcids.

Per poder incloure les dades genètiques en les anàlisis GLM, en principi es realitzaren arbres filogenètics amb les respectives seqüències emprant l'estratègia de la màxima parsimònia. Amb les seqüències de la regió *Core* es construïren dos arbres, un amb un valor de tall del 75% (*bootstrap majority rule*), en el qual les mostres s'agruparen segons el seu subtip, i un altre amb un valor de tall del 50%, en el qual es detectaren diferents agrupacions dins els clades de cada subtip, d'aquesta manera, aprofundim dins de la variabilitat de cada un dels subtip, sempre considerant variacions genètiques significatives. Amb les seqüències d'aminoàcids de la regió *NS5B* es construï un arbre amb un valor de tall del 50%.

La característica més remarcable d'aquests arbres és que en alguns genotips, les seqüències dels pacients s'agruparen en subclades augmentant així la resolució de l'anàlisi en front de si només s'empessin els genotips. En ambdós casos s'ha utilitzat una metodologia semblant, en cada cas, cada mostra dels clades filogenètics o dels genotips s'ha recodificat en una nova variable binària (0 absència, 1 presència) per sistematitzar el mètode. En el cas de les anàlisis *NS5B* també es transformaren cada subtip en variables binàries.

Per cada model, les variables de l'hoste i els clades filogenètics o genotípics foren emprats com a variables predictores potencials del grau de fibrosi en la nostra cohort de pacients coinfectats amb el VHC i el VIH.

Per evitar interferències amb el model final, es varen excloure mostres atípiques en l'anàlisi GLM (per exemple es va excloure la mostra 25LN-2a, degut a que era l'únic representant del genotip 2 en la nostra cohort), d'aquesta manera es volia evitar el biaix produït per mostres atípiques que s'havia detectat a l'estudi previ.

En primer lloc, les anàlisis GLM es varen dur a terme amb els tres índexs per separat i conjuntament, i seguint la metodologia desenvolupada per Szmargd *et al.* (2006), amb algunes modificacions. Per cada model, algunes variables de l'hoste i alguns clades filogenètics del VHC foren identificats com a predictors més importants i representatius de la fibrosi en la cohort de pacients coinfectats amb els VHC i el VIH.

En les anàlisis realitzades utilitzant les seqüències de la regió *Core*, el genotip 3a fou el que mostrà una variabilitat genètica més alta i una desviació estàndard més alta també en

els índexs de fibrosi. Això era degut a la presència d'un pacient (17LO3a) amb el recompte plaquetari més baix. En un primer moment, no es va excloure aquest pacients de les anàlisis preliminars, per veure quin efecte podria tenir aquest valor atípic.

En totes les anàlisis realitzades amb les seqüències de la regió *Core*, el rati GOT/GPT (o AST/ALT) presentava un gran pes estadísticament significatiu en els models seleccionats, donant, per tant, una explicació molt significativa als índexs de fibrosi, la qual cosa indicava la importància de les transaminases com a indicadors indirectes de la necrosi de cèl·lules hepàtiques (Gressner *et al.*, 2007). Es podria pensar que aquesta correlació era una conseqüència directa de la incorporació de les transaminases a les fórmules utilitzades per calcular els índexs APRI i FIB4, emperò aquests enzims no apareixen en el càlcul de l'índex IF, i segueixen resultant altament significatives quan es realitzaren les anàlisis emprant aquest índex. La mateixa explicació podria aplicar-se a l'enzim GGT, emprat en la fórmula de l'índex IF, però en aquest cas, el GGT no fou seleccionat en cap anàlisi basat en aquest índex. En el cas de l'edat, la presència com a variable predictiva en algunes de les anàlisis sembla ser conseqüència del procés estadístic i no de la inclusió en els càlculs d'alguns índexs.

La ràtio GOT/GPT va ser seleccionada en totes les anàlisis realitzades, possiblement perquè el pacient amb els tres índexs més elevats (17LO3a) també presentava un valor molt elevat per aquest paràmetre. De fet, les característiques d'aquest pacient ens permeten validar la metodologia GLM, ja que l'anàlisi detecta una associació negativa amb la ruta de transmissió del virus per l'ús de drogues parenterals, degut a que aquest pacient no fou infectat per aquesta via, sinó per un contagi heterosexual, encara que la significació és més baixa que en el cas de la ratio de transaminases. Per comprovar aquest *efecte d'atracció*, es varen realitzar les anàlisis sense aquest pacient, i els resultats foren bastant diferents. La ràtio GOT/GPT va continuar sent significant i seguí la mateixa tendència que abans; però la via de transmissió per ús de drogues parenterals va presentar una tendència positiva, sent directament relacionat amb el dany hepàtic. Al mateix temps, algunes noves variables aparegueren com a significatives, com per exemple el tractament antiretroviral, encara que gairebé tots els pacients foren tractats amb el mateix tractament. Aquest resultats suggereixen un efecte significatiu d'aquest tractament a l'hora d'explicar el dany hepàtic, descrit per altres autors a la bibliografia (Matthews i Dore, 2008). En totes les anàlisis amb el pacient atípic inclòs, el clúster del VHC més seleccionat fou el relacionat amb el genotip

3a del virus. Sense ell, els clústers seleccionats són més diversos, indicant la forta influència d'aquest pacient. Entre els clústers, tenen importància els relacionats amb els genotips 1a i 1b, més en concordança amb la resposta dels pacients al dany hepàtic.

En les anàlisis realitzades amb les seqüències de la regió *NS5B*, es va intentar millorar i perfeccionar la metodologia de l'anàlisi GLM. Per això, ja no es tengueren directament en compte les mostres atípiques observades en les anàlisis anteriors. A més, per tal d'evitar els possibles biaixos (encara que mínims) que es podrien derivar de la inclusió de determinades variables en els índexs, no s'utilitzaren com a variables predictives les incloses en els càlculs dels índexs de fibrosi: GOT, GPT, edat i recompte plaquetari. D'aquesta manera, en l'anàlisi basat en els clades filogenètics i les mesures de fibrosi independents, es va trobar que l'enzim GGT presentava una forta relació estadísticament positiva amb la fibrosi, mentre que el fet de no presentar l'estadi C3 de la malaltia causada pel VIH i els nivells de colesterol estaven estadísticament associats, però de forma negativa amb la fibrosi. Aquests resultats es troben en concordança amb els resultats clínics trobats en altres estudis. Així, Kau *et al.* (2008) detectaren nivells elevats de GGT en pacients amb esteatosi hepàtics i fibrosi avançada. A més, la classificació de la infecció per VIH en l'estadi C3 és un indicador de que la malaltia ha arribat a l'estadi de SIDA, al mateix temps que es presenta un recompte de cèl·lules T CD4+ extremadament baix. Alguns estudis suggereixen que un estadi avançat de la infecció pel VIH condueix a un augment del risc de malaltia hepàtica en pacients coinfectats. Per tant, estadis primerencs de la infecció pel VIH (estadi no C3) serien consistent amb uns nivells baixos de fibrosis (Matthews i Dore, 2008). Finalment, un estudi més recent (Piver *et al.*, 2010), ha mostrat que durant el transcurs de la infecció pel VHC, el metabolisme lipídic es veu altament alterat, mostrant que en pacients que han desenvolupat hepatitis C crònica (i per tant, més susceptibles a patir fibrosi hepàtica), presenten nivells de triglicèrids i colesterol més baixos que els pacients que foren infectats amb el virus però l'eliminaren espontàniament. Això explicaria la relació negativa entre els nivells de colesterol i el dany hepàtic. No obstant això, la majoria dels estudis en aquest camp s'han dut a terme en cohorts de pacients monoinfectats, i els resultats del present estudi estan mostrant també aquestes relacions en individus coinfectats VHC/VIH (Piver *et al.*, 2010).

Trobar característiques genètiques de les seqüències del VHC potencialment associades amb fibrosi, potser clínicament rellevant. En el darrer estudi, el clade F

corresponent als pacients del subtip 3a, fou positivament associat a la fibrosi hepàtica pel GLM, mentre el clade E (pacients 3LN-3a i 12LN-3a), corresponent a un subclade dels pacients de genotip 3a presentà una associació negativa amb la fibrosi. Cap clade fou seleccionat en les anàlisis utilitzant només les dades dels genotips, ni tampoc quan s'utilitzaren els índexs de fibrosi corregits pel temps estimat d'infecció com a variable resposta. Aquests resultats indiquen que considerar la variabilitat intragenotípica permet fer associacions més enllà del genotip o subtip viral, possibilitant l'augment de la resolució o especificitat de les anàlisis. S'ha de destacar que la variabilitat genètica de la regió *NS5B* (i el mateix passava en la regió *Core*) del VHC fou més elevada en els virus pertanyents al genotip 3a, genotip que incloïa un subclade diferenciat de la resta de mostres d'aquest genotip.

La importància del mètode GLM és patent quan es comparen les anàlisis resultants emprant només les dades dels genotips, amb les resultants de les anàlisis de les dades procedents de l'arbre filogenètic. Mentre que en el primer cas, l'enzim GGT fou la única variable seleccionada (i cap genotip/subtip fou seleccionat), emprant les dades filogenètiques, com s'ha indicat anteriorment, alguns subclades de l'arbre també foren seleccionats com a variables predictives. Aquest fet posa èmfasis en la rellevància de les anàlisis emprant la diversitat genètica del virus basada en seqüències nucleotídiques o aminoacídiques i no només basada en la determinació del genotip/subtip del virus.

També es realitzaren anàlisis GLM utilitzant els índexs de fibrosi corregits pel temps estimat d'infecció amb la finalitat de tenir en compte d'alguna manera la taxa de progressió de fibrosi. En aquestes anàlisis, es varen obtenir els mateixos resultats utilitzant les dades procedents dels genotips i de l'arbre filogenètic. En cap cas no es va seleccionar cap genotip ni cap clúster en cap model, l'enzim GGT fou seleccionat amb una tendència positiva, i el colesterol amb una tendència negativa. D'aquesta manera, aquests resultats indicarien que aquestes dues variables serien les variables candidates més significatives a estar relacionades amb el desenvolupament de la fibrosi en la nostra cohort. De totes maneres, s'ha de notar que, malgrat la taxa de desenvolupament de fibrosi pugui resultar més rellevant que una mesura puntual, els índexs emprats estan descrits com a mesures puntuals de fibrosi. Per tant, la correcció feta pel temps d'infecció podria distorsionar els índexs; així, els resultats obtinguts emprant aquesta correcció s'han d'interpretar amb precaució.

Entre les possibles explicacions de perquè dos clústers de pacients infectats amb el genotip 3a presenten diferents nivells de fibrosi, es podria especular que podria haver diferències a nivell d'aminoàcids que podrien determinar aquests canvis. Analitzant a fons les seqüències, es pot observar que els dos clades difereixen en un sol aminoàcid (posició 273 en la proteïna *NS5B* i 2793 en la poliproteïna). En les mostres pertanyents al clade E, aquesta posició és ocupada per una prolina, mentre que en la resta de mostres del genotip 3a, aquesta posició està ocupada per una glutamina, emperò l'efecte que pot tenir aquest canvi en la replicació viral roman desconegut.

2. ESTUDI DELS POLIMORFISMES EN PACIENTS versus RESPOSTA AL TRACTAMENT

S'ha establert que la resposta dels pacients al tractament a la infecció pel VHC emprant interferó pegilat i ribavirina depèn de factors del virus i de l'hoste (Rauch *et al.*, 2010a). En relació amb els factors relacionats amb l'hoste, en el darrer estudi que s'ha dut a terme en aquesta tesi es va estudiar el paper de diferents polimorfismes: IL28b, CTLA4 (+49 i -318), LDLr, polimorfismes en els gen HFE (C282Y, H63D i S65C) i l'haplogrup del DNA mitocondrial. L'objectiu era conèixer fins a quin punt, aquests polimorfismes genètics poguessin afectar la resposta al tractament en una cohort de pacients coinfectats procedent de l'hospital Son Llàtzer de Palma, Mallorca.

La freqüència del polimorfisme IL28b de la nostra cohort va ser similar a la trobada en la cohort europea-americana analitzada a Ge *et al.* (2009), encara que la freqüència del genotip CC va ser més baixa que la trobada en altres cohorts espanyoles (Montes-Cano *et al.*, 2010; Rallón *et al.*, 2010). A més, no es varen trobar diferències entre els pacients i els controls. D'una forma similar, les freqüències dels polimorfismes CTLA4 i LDLr foren semblants a les trobades en altres poblacions (Yee *et al.*, 2003; Mas Marques *et al.*, 2009). No obstant això, en la mutació H63D del gen HFE es trobaren diferències significatives entre la població control analitzada a un estudi anterior (Picornell *et al.*, 2005) i la cohort de pacients coinfectats d'aquest estudi. S'ha de destacar que Espanya és un dels països del món amb una freqüència més elevada de dita mutació (Guix *et al.*, 2002). La distribució de freqüències de la resta de polimorfismes i dels haplogrups del DNA mitocondrial (mtDNA) no mostraren cap diferència entre el grup de pacients i el de controls.

La mitjana de resposta al tractament de la cohort estudiada és similar a la trobada a la literatura, encara que la ràtio d'aclariment espontani fou una mica menor (Seeff, 2009).

Tal com també queda reflectit en la bibliografia, les anàlisis relacionant el genotip del virus que infectava els pacients i la seva resposta viral sostinguda (RVS), mostraren que els pacients infectats amb els genotips 1 i 4 estaven relacionats amb una pitjor resposta al tractament, mentre els infectats amb el genotip 3 responien millor (Rauch *et al.*, 2010a).

La pregunta més important del present estudi va ser trobar alguna relació entre el polimorfisme genètic que presenta l'hoste i la resposta al tractament. El polimorfisme IL28B ha estat el més útil, juntament amb l'ús tradicional dels genotips virals, per anticipar la resposta d'un pacient al tractament. En la nostra cohort de pacients, el genotip CC del polimorfisme IL28b es troba molt poc representat entre els pacients infectats amb el genotip 4. Els pacients CC responen molt bé al tractament, mentre que els pacients infectats amb el genotip 4 estan entre els que responen pitjor, per tal cosa, és difícil entreveure quin dels dos efectes té més importància. Aquestes interaccions s'han de tenir en compte per determinar si un polimorfisme dona un millor pronòstic per la resposta al tractament. Les dades més significants referents al polimorfisme IL28b, les trobem en els pacients portadors del genotip CC en la resposta a les 4 setmanes d'haver començament el tractament, per tant es podria deduir que el genotip CC afavoreix una ràpida resposta al tractament. Els resultats són més significatius quan es comparen els pacients portadors del genotip CC, en front del grup format pels pacients heterozigots i els homozigots TT. Més encara, s'ha trobat que només un pacient amb el genotip CC infectat amb un virus del genotip 2, no ha assolit la RVS. Mentre que només un pacients amb el genotip TT i infectat amb el genotip 1 o 4 del virus ha assolit la RVS.

L'anàlisi dels pacients que havien recaigut, és a dir, que en un determinat moment sí que respongueren al tractament, malgrat no assoliren la RVS, mostrà que només un dels pacients que mai havia respost al tractament presentava el genotip CC. Això podria indicar la importància d'aquest polimorfisme en una resposta primerenca al tractament, més que un efecte damunt la RVS.

No es trobà cap efecte entre el polimorfisme LDLr i la resposta al tractament. En algunes publicacions es pot veure que el genotip GG mostra una millor resposta (Hennig *et al.*, 2002; Pineda *et al.*, 2011), però en la nostra cohort només s'ha trobat una relació entre el genotip AA i la presència d'esteatosi hepàtica.

L'aclariment espontani es va trobar que era més elevat en pacients amb l'al·lel G en el polimorfisme CTLA4 +49. Analitzant els motius d'aquest resultat es va veure que els 6

individus que presentaven aclariment espontani, eren portadors d'aquest al·lel. Aquest resultat està en concordança amb les dades trobades a Yee *et al.* (2003), on es trobà que aquest al·lel millorava la resposta al tractament en pacients monoinfectats amb el genotip 1 i tractats amb interferó no pegilat. Altres estudis (Nischalke *et al.*, 2010) concloueren que el genotip GG estava associat amb una millor resposta al tractament i no amb l'aclariment espontani com s'ha trobat a la nostra cohort.

Les mutacions en el gen HFE pareix que no tenen cap connexió amb les variables estudiades. Malgrat això, s'ha de destacar que tots els pacients homozigots H63D, tots els heterozigots S65C i gairebé tots els heterozigots C282Y foren sotmesos a l'obtenció de biòpsies hepàtiques. Retrospectivament, aquest fet podria conduir a pensar que aquests pacients, en algun moment, presentaren alguna evidència objectiva per haver estat sotmesos a aquest procediment, encara que no es reflectís en els nivells de fibrosi o d'altres variables.

En conclusió, en aquest estudi s'ha volgut destacar la importància de conèixer els antecedents genètics dels pacients abans de començar un tractament, tenint en compte les característiques específiques de la cohort. Els resultats han mostrat la importància del genotip CC en el polimorfisme d'IL28B alhora d'explicar la predisposició a una millor resposta al tractament pegIFN-RIB, i també, una major susceptibilitat a l'aclariment viral espontani en portadors de l'al·lel G en polimorfismes de CTLA4 +49.



CONCLUSIONS

1. La metodologia GLM permet la combinació d'eines filogenètiques amb el modelatge estadístic, integrant, en una única anàlisi, diferent tipus de variables de l'hoste (demogràfiques, epidemiològiques i clíniques) amb dades genètiques del genoma viral en una única anàlisi integrada, seleccionant els factors que millor expliquen la variable de resposta.
2. L'ús dels marcadors indirectes de la fibrosi hepàtica (dany hepàtic) presenta limitacions, malgrat això, són útils en estudis retrospectius en els quals manquen dades de biòpsies.
3. La presència d'un pacient amb dades atípiques (outlier) influeix en la selecció d'algunes variables en els models. La inclusió i exclusió d'aquest pacient en les anàlisis de GLM ens permet validar millor els models seleccionats i contrastar les variables seleccionades pels diferents models.
4. Alguns clades filogenètics del virus de l'hepatitis C estan relacionats amb valors elevats dels marcadors de la fibrosi (genotip 3a). La presència de subclades dins els genotips, i la seva inclusió en els models, augmenta la resolució de les anàlisis.
5. Depenent dels diferents models, les variables seleccionades de forma positiva han estat: la utilització del tractament antiretroviral, l'administració de drogues via parenteral i els nivells de l'enzim GGT; mentre que entre les seleccionades negativament, es troba el tenir un nivell baix de càrrega viral del virus HIV, el no trobar-se en un estadi de la malaltia per VIH C3 i els nivells de colesterol.
6. Les freqüències gèniques i genotípiques dels polimorfismes estudiats no difereixen entre població control i cohort de pacients, excepte en el cas de la mutació H63D del gen HFE.
7. Els pacients infectats amb virus dels genotips 1 o 4 responen pitjor al tractament amb interferó pegilat i ribavirina que els pacients infectats amb virus del genotip 3.
8. La freqüència del genotip CC del polimorfisme IL28b en la població estudiada és més baixa que en altres cohorts espanyoles. Aquest genotip gairebé no està representat en la nostra cohort entre els pacients infectats amb el genotip 4 del virus.
9. El genotip CC del polimorfisme IL28b s'ha relacionat amb una major resposta viral sostinguda, però més significativament amb una resposta ràpida al tractament (a les 4 setmanes). Només un pacient amb el genotip CC no ha respost en cap moment al

tractament. Aquest efecte és més evident, quan es comparen els pacients amb el genotip CC amb el grup de pacients formats pels que tenen genotip CT i TT.

10. Els sis pacients que presentaren aclariment espontani de la malaltia són portadors de l'al·lel G en el polimorfisme CTLA4 +49.
 11. Retrospectivament s'ha trobat que tots els pacients homozigots H63D, heterozigots S65C i gairebé tots els heterozigots C282Y foren sotmesos a l'obtenció de biòpsia.
 12. El polimorfisme IL28b (rs12979860) ha estat el més útil, juntament amb els genotips virals per tal de preveure la resposta d'un pacient coinfectats amb els virus de l'hepatitis C i el virus de la immunodeficiència humana al tractament amb interferó pegilat i ribavirina.
-

- (1993). From the Centers for Disease Control and Prevention. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *JAMA* 269 (6): 729-730.
- Ahmad W, Ijaz B, Gull S, Asad S, Khaliq S, Jahan S, Sarwar MT, Kausar H, *et al.* (2011). A brief review on molecular, genetic and imaging techniques for HCV fibrosis evaluation. *Virol J* 8 (1): 53.
- Alter H (2006a). Viral hepatitis. *Hepatology* 43 (2 Suppl 1): S230-234.
- Alter MJ (2006b). Epidemiology of viral hepatitis and HIV co-infection. *J Hepatol* 44 (1 Suppl): S6-9.
- Alter MJ (2007). Epidemiology of hepatitis C virus infection. *World J Gastroenterol* 13 (17): 2436-2441.
- Asselah T, Bieche I, Sabbagh A, Bedossa P, Moreau R, Valla D, Vidaud M i Marcellin P (2009). Gene expression and hepatitis C virus infection. *Gut* 58 (6): 846-858.
- Asselah T, Estrabaud E, Bieche I, Lapalus M, De Muynck S, Vidaud M, Saadoun D, Soumelis V, *et al.* (2010). Hepatitis C: viral and host factors associated with non-response to pegylated interferon plus ribavirin. *Liver Int* 30 (9): 1259-1269.
- Beaulieu PL (2009). Recent advances in the development of NS5B polymerase inhibitors for the treatment of hepatitis C virus infection. *Expert opinion on therapeutic patents* 19 (2): 145-164.
- Bedossa P, Dargere D i Paradis V (2003). Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 38 (6): 1449-1457.
- Bonkovsky HL, Naishadham D, Lambrecht RW, Chung RT, Hoefs JC, Nash SR, Rogers TE, Banner BF, *et al.* (2006). Roles of iron and HFE mutations on severity and response to therapy during retreatment of advanced chronic hepatitis C. *Gastroenterology* 131 (5): 1440-1451.
- Bonkovsky HL, Troy N, McNeal K, Banner BF, Sharma A, Obando J, Mehta S, Koff RS, *et al.* (2002). Iron and HFE or TfR1 mutations as comorbid factors for development and progression of chronic hepatitis C. *J Hepatol* 37 (6): 848-854.
- Bruguera M i Forn X (2006). Hepatitis C en España. *Med Clin (Barc)* 127 (3): 113-117.
- Buti M i Esteban R (2003). VII International Symposium on Viral hepatitis. Madrid, Acción Médica.
- Carneiro MV, Souza FF, Teixeira AC, Figueiredo JF, Villanova MG, Secaf M, Passos AD, Ramalho LN, *et al.* (2010). The H63D genetic variant of the HFE gene is independently associated with the virological response to interferon and

- ribavirin therapy in chronic hepatitis C. *Eur J Gastroenterol Hepatol* 22 (10): 1204-1210.
- Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW i Houghton M (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244 (4902): 359-362.
- Cifuentes C, Juan M, Sastre M, Payeras A, Fernández-Baca V, Homar F, Gallegos C i Bassa A (2004). Coinfected by HIV-HCV: epidemiological data. First genotype determination in a sample from Mallorca Island. 27th World Congress of Internal Medicine, Granada, Spain.
- Crawley MJ (2007). *The R book*. Chichester, England, Hoboken, N.J., Wiley.
- Dubuisson J (2007). Hepatitis C virus proteins. *World J Gastroenterol* 13 (17): 2406-2415.
- Echevarría JM, León P, Pozo F i Avellón A (2006). Follow-up of the prevalence of hepatitis C virus genotypes in Spain during a nine-year period (1996-2004). *Enferm Infecc Microbiol Clin* 24 (1): 20-25.
- Eisenbach C, Gehrke SG i Stremmel W (2004). Iron, the HFE gene, and hepatitis C. *Clin Liver Dis* 8 (4): 775-785, vii-viii.
- Esteban JI, Sauleda S i Quer J (2008). The changing epidemiology of hepatitis C virus infection in Europe. *J Hepatol* 48 (1): 148-162.
- Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, Domingo R, Jr., *et al.* (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 13 (4): 399-408.
- Feinstone SM, Kapikian AZ, Purcell RH, Alter HJ i Holland PV (1975). Transfusion-associated hepatitis not due to viral hepatitis type A or B. *New Engl J Med* 292 (15): 767-770.
- Forns X, Ampurdanès S, Llovet JM, Aponte J, Quinto L, Martínez-Bauer E, Bruguera M, Sánchez-Tapias JM, *et al.* (2002). Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology* 36 (4 Pt 1): 986-992.
- Gane E (2011). Future hepatitis C virus treatment: interferon-sparing combinations. *Liver Int* 31 (Suppl 1): 62-67.
- García-Álvarez M, Guzmán-Fulgencio M, Berenguer J, Micheloud D, Campos Y, López JC, Cosín J, Miralles P, *et al.* (2011). European mitochondrial DNA haplogroups and liver fibrosis in HIV and hepatitis C virus coinfecting patients. *AIDS* 25 (13): 1619-1926.

- Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, *et al.* (2009). Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461 (7262): 399-401.
- Geier A, Reugels M, Weiskirchen R, Wasmuth HE, Dietrich CG, Siewert E, Gartung C, Lorenzen J, *et al.* (2004). Common heterozygous hemochromatosis gene mutations are risk factors for inflammation and fibrosis in chronic hepatitis C. *Liver Int* 24 (4): 285-294.
- Germani G, Burroughs AK i Dhillon AP (2010). The relationship between liver disease stage and liver fibrosis: a tangled web. *Histopathology* 57 (6): 773-784.
- Gigou M, Roque-Afonso AM, Falissard B, Penin F, Dussaix E i Feray C (2001). Genetic clustering of hepatitis C virus strains and severity of recurrent hepatitis after liver transplantation. *J Virol* 75 (23): 11292-11297.
- González-Candelas F, Bracho MA i Moya A (2003). Molecular epidemiology and forensic genetics: application to a hepatitis C virus transmission event at a hemodialysis unit. *J Infect Dis* 187 (3): 352-358.
- Gressner OA, Weiskirchen R i Gressner AM (2007). Biomarkers of hepatic fibrosis, fibrogenesis and genetic pre-disposition pending between fiction and reality. *J Cell Mol Med* 11 (5): 1031-1051.
- Guix P, Picornell A, Parera M, Galmes A, Obrador A, Ramon MM i Castro JA (2002). Distribution of HFE C282Y and H63D mutations in the Balearic Islands (NE Spain). *Clin Genet* 61 (1): 43-48.
- Hall TA (1999). Bioedit: a user friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95-98.
- Hanson EH, Imperatore G i Burke W (2001). HFE gene and hereditary hemochromatosis: a HuGE review. *Human Genome Epidemiology. Am J Epidemiol* 154 (3): 193-206.
- He Y i Tan S-L (2011). *Hepatitis C : Antiviral Drug Discovery and Development*. Norfolk, UK, Caister Academic Press.
- Heathcote EJ (2007). Antiviral therapy: chronic hepatitis C. *J Viral Hepatitis* 14 (Suppl 1): 82-88.
- Hendrickson SL, Hutcheson HB, Ruiz-Pesini E, Poole JC, Lautenberger J, Sezgin E, Kingsley L, Goedert JJ, *et al.* (2008). Mitochondrial DNA haplogroups influence AIDS progression. *AIDS* 22 (18): 2429-2439.
- Hendrickson SL, Kingsley LA, Ruiz-Pesini E, Poole JC, Jacobson LP, Palella FJ, Bream JH, Wallace DC, *et al.* (2009). Mitochondrial DNA haplogroups influence

- lipoatrophy after highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* 51 (2): 111-116.
- Hennig BJ, Hellier S, Frodsham AJ, Zhang L, Klenerman P, Knapp S, Wright M, Thomas HC, *et al.* (2002). Association of low-density lipoprotein receptor polymorphisms and outcome of hepatitis C infection. *Genes Immun* 3 (6): 359-367.
- Ishizu Y, Katano Y, Honda T, Hayashi K, Ishigami M, Itoh A, Hirooka Y, Nakano I, Goto H (2012). Clinical impact of HFE mutations in Japanese patients with chronic hepatitis C. *J Gastroenterol Hepatol.* 27(6): 1112-1116.
- Jiménez-Hernaández N, Torres-Puente M, Bracho MA, García-Robles I, Ortega E, del Olmo J, Carnicer F, González-Candelas F, *et al.* (2007). Epidemic dynamics of two coexisting hepatitis C virus subtypes. *J Gen Virol* 88 (Pt 1): 123-133.
- Kau A, Vermehren J i Sarrazin C (2008). Treatment predictors of a sustained virologic response in hepatitis B and C. *J Hepatol* 49 (4): 634-651.
- Khan M, Jahan S, Khaliq S, Ijaz B, Ahmad W, Samreen B i Hassan S (2010). Interaction of the hepatitis C virus (HCV) core with cellular genes in the development of HCV-induced steatosis. *Arch Virol* 155 (11): 1735-1753.
- Kloss-Brandstaetter A, Pacher D, Schoenherr S, Weissensteiner H, Binna R, Specht G, Kronenberg F (2010). HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups.
<http://www.haplogrep.uibk.ac.at> doi: 10.1002/humu.21382
- Kolaczkowski B i Thornton JW (2004). Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* 431 (7011): 980-984.
- Kuiken C, Hraber P, Thurmond J i Yusim K (2008). The hepatitis C sequence database in Los Alamos. *Nucleic Acids Res* 36 (Database issue): D512-516.
- Le Gac G i Ferec C (2005). The molecular genetics of haemochromatosis. *Eur J Hum Genet* 13 (11): 1172-1185.
- Li JH, Lao XQ, Tillmann HL, Rowell J, Patel K, Thompson A, Suchindran S, Muir AJ, *et al.* (2010). Interferon-lambda genotype and low serum low-density lipoprotein cholesterol levels in patients with chronic hepatitis C infection. *Hepatology* 51 (6): 1904-1911.
- Librado P i Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25 (11): 1451-1452.

- Lindenbach BD i Rice CM (2005). Unravelling hepatitis C virus replication from genome to function. *Nature* 436 (7053): 933-938.
- López-Labrador FX, Ampurdanès S, Fornés X, Castells A, Sáiz JC, Costa J, Bruix J, Sánchez Tapias JM, *et al.* (1997). Hepatitis C virus (HCV) genotypes in Spanish patients with HCV infection: relationship between HCV genotype 1b, cirrhosis and hepatocellular carcinoma. *J Hepatol* 27 (6): 959-965.
- López-Labrador FX, Berenguer M, Sempere A, Prieto M, Sirera R, González-Molina A, Ortiz V, Marty ML, *et al.* (2004). Genetic variability of hepatitis C virus NS3 protein in human leukocyte antigen-A2 liver transplant recipients with recurrent hepatitis C. *Liver Transpl* 10 (2): 217-227.
- López-Labrador FX, Bracho MA, Berenguer M, Coscolla M, Rayon JM, Prieto M, Carrasco D, Gómez MD, *et al.* (2006). Genetic similarity of hepatitis C virus and fibrosis progression in chronic and recurrent infection after liver transplantation. *J Viral Hepatitis* 13 (2): 104-115.
- López-Labrador FX, Dove L, Hui CK, Phung Y, Kim M, Berenguer M i Wright TL (2007). Trends for genetic variation of Hepatitis C Virus quasispecies in Human Immunodeficiency virus-1 coinfecting patients. *Virus Res* 130 (1-2): 285-291.
- Lynas C (1997). A cheaper and more rapid polymerase chain reaction-restriction fragment length polymorphism method for the detection of the HLA-H gene mutations occurring in hereditary hemochromatosis. *Blood* 90 (10): 4235-4236.
- Macías J, Berenguer J, Japón MA, Girón JA, Rivero A, López-Cortés LF, Moreno A, González-Serrano M, *et al.* (2009). Fast fibrosis progression between repeated liver biopsies in patients coinfecting with human immunodeficiency virus/hepatitis C virus. *Hepatology* 50 (4): 1056-1063.
- Mas Marques A, Mueller T, Welke J, Taube S, Sarrazin C, Wiese M, Halangk J, Witt H, *et al.* (2009). Low-density lipoprotein receptor variants are associated with spontaneous and treatment-induced recovery from hepatitis C virus infection. *Infect Genet Evol* 9 (5): 847-852.
- Matthews GV i Dore GJ (2008). HIV and hepatitis C coinfection. *J Gastroenterol Hepatol* 23 (7 Pt 1): 1000-1008.
- McCaughan GW i George J (2004). Fibrosis progression in chronic hepatitis C virus infection. *Gut* 53 (3): 318-321.
- Micheloud D, Berenguer J, Guzmán-Fulgencio M, Campos Y, García-Álvarez M, Catalán P, Cosín J, Miralles P, *et al.* (2011). European mitochondrial DNA haplogroups and metabolic disorders in HIV/HCV-coinfecting patients on highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* 58 (4): 371-378.

- Molina S, Castet V, Fournier-Wirth C, Pichard-Garcia L, Avner R, Harats D, Roitelman J, Barbaras R, *et al.* (2007). The low-density lipoprotein receptor plays a role in the infection of primary human hepatocytes by hepatitis C virus. *J Hepatol* 46 (3): 411-419.
- Montes-Cano MA, García-Lozano JR, Abad-Molina C, Romero-Gomez M, Barroso N, Aguilar-Reina J, Nunez-Roldan A i Gonzalez-Escribano MF (2010). Interleukin-28B genetic variants and hepatitis virus infection by different viral genotypes. *Hepatology* 52 (1): 33-37.
- Nischalke HD, Vogel M, Mauss S, Baumgarten A, Lutz T, Danta M, Naumann U, Coenen M, *et al.* (2010). The cytotoxic lymphocyte antigen 4 polymorphisms affect response to hepatitis C virus-specific therapy in HIV(+) patients with acute and chronic hepatitis C virus co-infection. *AIDS* 24 (13): 2001-2007.
- Njajou OT, Alizadeh BZ i van Duijn CM (2004). Is genetic screening for hemochromatosis worthwhile? *Eur J Epidemiol* 19 (2): 101-108.
- Ohno O, Mizokami M, Wu RR, Saleh MG, Ohba K, Orito E, Mukaide M, Williams R, *et al.* (1997). New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. *J Clin Microbiol* 35 (1): 201-207.
- Penin F, Dubuisson J, Rey FA, Moradpour D i Pawlotsky JM (2004). Structural biology of hepatitis C virus. *Hepatology* 39 (1): 5-19.
- Picornell A, Gomez-Barbeito L, Tomas C, Castro JA i Ramon MM (2005). Mitochondrial DNA HVRI variation in Balearic populations. *Am J Phys Anthropol* 128 (1): 119-130.
- Pineda JA, Caruz A, Di Lello FA, Camacho A, Mesa P, Neukam K, Rivero-juárez A, Macías J, *et al.* (2011). Low-density lipoprotein receptor genotyping enhances the predictive value of IL28B genotype in HIV/hepatitis C virus-coinfected patients. *AIDS* 25 (11): 1415-1420.
- Piver E, Roingard P i Pagès JC (2010). The cell biology of hepatitis C virus (HCV) lipid addiction: molecular mechanisms and its potential importance in the clinic. *Int J Biochem Cell Biol* 42 (6): 869-879.
- Poynard T, Yuen MF, Ratziu V i Lai CL (2003). Viral hepatitis C. *Lancet* 362 (9401): 2095-2100.
- Pybus OG, Cochrane A, Holmes EC i Simmonds P (2005). The hepatitis C virus epidemic among injecting drug users. *Infect Genet Evol* 5 (2): 131-139.
- Rallón NI, Naggie S, Benito JM, Medrano J, Restrepo C, Goldstein D, Shianna KV, Vispo E, *et al.* (2010). Association of a single nucleotide polymorphism near the

- interleukin-28B gene with response to hepatitis C therapy in HIV/hepatitis C virus-coinfected patients. *AIDS* 24 (8): F23-29.
- Ramos Paesa C, Marcilla F, López G, Hueso E, Pascual A i Aguirre JM (2007). Valuation of APRI and Forns models for non-invasive diagnosis of fibrosis in patients with hepatitis C in coinfecting and non-coinfecting with HIV. *An Med Interna* 24 (8): 369-374.
- Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, Bochud M, Battegay M, *et al.* (2010a). Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* 138 (4): 1338-1345, 1345 e1331-1337.
- Rauch A, Rohrbach J i Bochud PY (2010b). The recent breakthroughs in the understanding of host genomics in hepatitis C. *Eur J Clin Invest* 40 (10): 950-959.
- Reiberger T, Ferlitsch A, Sieghart W, Kreil A, Breitenecker F, Rieger A, Schmied B, Gangl A, *et al.* (2010). HIV-HCV co-infected patients with low CD4+ cell nadirs are at risk for faster fibrosis progression and portal hypertension. *J Viral Hepatitis* 17 (6): 400-409.
- Rivero-Juárez A, Camacho A, Caruz A, Neukam K, Gonzalez R, Di Lello FA, Perez-Camacho I, Mesa P, *et al.* (2012). LDLr genotype modifies the impact of IL28B on HCV viral kinetics after the first weeks of treatment with PEG-IFN/RBV in HIV/HCV patients. *AIDS* 26 (8): 1009-1015.
- Rotman Y i Liang TJ (2009). Coinfection with hepatitis C virus and human immunodeficiency virus: virological, immunological, and clinical outcomes. *J Virol* 83 (15): 7366-7374.
- Schott E, Witt H, Hinrichsen H, Neumann K, Weich V, Bergk A, Halangk J, Muller T, *et al.* (2007). Gender-dependent association of CTLA4 polymorphisms with resolution of hepatitis C virus infection. *J Hepatol* 46 (3): 372-380.
- Seeff LB (2009). The history of the "natural history" of hepatitis C (1968-2009). *Liver Int* 29 (Suppl 1): 89-99.
- Shuhart MC, Sullivan DG, Bekele K, Harrington RD, Kitahata MM, Mathisen TL, Thomassen LV, Emerson SS, *et al.* (2006). HIV infection and antiretroviral therapy: effect on hepatitis C virus quasispecies variability. *J Infect Dis* 193 (9): 1211-1218.
- Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, Feinstone S, Halfon P, Inchauspe G, *et al.* (2005). Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* 42 (4): 962-973.

- Standish RA, Cholongitas E, Dhillon A, Burroughs AK i Dhillon AP (2006). An appraisal of the histopathological assessment of liver fibrosis. *Gut* 55 (4): 569-578.
- Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, M SS, Torriani FJ, *et al.* (2006). Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43 (6): 1317-1325.
- Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, *et al.* (2009). IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 41 (10): 1100-1104.
- Suppiah V, Gaudieri S, Armstrong NJ, O'Connor KS, Berg T, Weltman M, Abate ML, Spengler U, *et al.* (2011). IL28B, HLA-C, and KIR variants additively predict response to therapy in chronic hepatitis C virus infection in a European Cohort: a cross-sectional study. *PLoS medicine* 8 (9): e1001092.
- Szmaragd C, Nichols RA i Balloux F (2006). A novel approach to characterise pathogen candidate genetic polymorphisms involved in clinical outcome. *Infect Genet Evol* 6 (1): 38-45.
- Tamura K, Dudley J, Nei M i Kumar S (2007). MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24 (8): 1596-1599.
- Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, *et al.* (2009). Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 41 (10): 1105-1109.
- Theise ND (2007). Liver biopsy assessment in chronic viral hepatitis: a personal, practical approach. *Mod Pathol* 20 (Suppl 1): S3-14.
- Thenie A, Orhant M, Gicquel I, Fergelot P, Le Gall JY, David V i Mosser J (2000). The HFE gene undergoes alternate splicing processes. *Blood Cells Mol Dis* 26 (2): 155-162.
- Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, Kidd J, Kidd K, *et al.* (2009). Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* 461 (7265): 798-801.
- Thomson BJ i Finch RG (2005). Hepatitis C virus infection. *Clin Microbiol Infect* 11 (2): 86-94.
- Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, Dhalluin-Venier V, Fontaine H i Pol S (2007). FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. Comparison with liver biopsy and fibrotest. *Hepatology* 46 (1): 32-36.

- Venables WN i Ripley BD (1999). Generalized Linear Models. Modern applied statistics with S. New York, Springer: 183-210.
- Vertuani S, Bazzaro M, Gualandi G, Micheletti F, Marastoni M, Fortini C, Canella A, Marino M, *et al.* (2002). Effect of interferon-alpha therapy on epitope-specific cytotoxic T lymphocyte responses in hepatitis C virus-infected individuals. *Eur J Immunol* 32 (1): 144-154.
- Vigilant L, Pennington R, Harpending H, Kocher TD i Wilson AC (1989). Mitochondrial DNA sequences in single hairs from a southern African population. *Proc Natl Acad Sci U S A* 86 (23): 9350-9354.
- Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS i Lok AS (2003). A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 38 (2): 518-526.
- Wallace DC (2005). A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annu Rev Genet* 39: 359-407.
- Ye J (2007). Reliance of host cholesterol metabolic pathways for the life cycle of hepatitis C virus. *PLoS Pathog* 3 (8): e108
- Yee LJ, Perez KA, Tang J, van Leeuwen DJ i Kaslow RA (2003). Association of CTLA4 polymorphisms with sustained response to interferon and ribavirin therapy for chronic hepatitis C virus infection. *J Infect Dis* 187 (8): 1264-1271.

BIBLIOGRAFIA PERSONAL

Publicacions:

- Matas M, Guix P, Castro JA, Parera M, Ramon MM, Obrador A i Picornell A (2006). Prevalence of HFE C282Y and H63D in Jewish populations and clinical implications of H63D homozygosity. *Clin Genet* 69 (2): 155-162.
- Matas M, Picornell A, Cifuentes C, Payeras A, Bassa A, Homar F, López-Labrador FX, Moya A, Ramon MM, Castro JA (2010). Relating the liver damage with hepatitis C virus polymorphism in core region and human variables in HIV-1-coinfected patients. *Infect Genet Evol* 10, 1252-1261.
- Matas M, Picornell A, Cifuentes C, Payeras A, Bassa A, Homar F, González-Candelas F, López-Labrador FX, Moya A, Ramon MM, Castro JA (2012). Generalized Linear Model (GLM) framework for the association of host variables and viral strains with liver fibrosis in HCV/HIV coinfecting patients. *In Press Infect Genet Evol* (Oct 2012).
- Matas M, Picornell A, Cifuentes C, Payeras A, Bassa A, Homar F, González-Candelas F, López-Labrador FX, Moya A, Ramon MM, Castro JA (2012). Relating the outcome of HCV infection with different host SNP polymorphisms in a Majorcan coinfecting HCV-HIV population treated with pegIFN-RIB. Manuscript in preparation.

Comunicacions a congressos:

- Matas M, Castro JA, Ramon MM, Guix P, Parera M, Obrador A, Picornell A. Hereditary hemochromatosis mutations C282Y and H63D in non-ashkenazi Jewish populations. *European Iron Club 2006* (Barcelona, Spain). 28-29 setembre 2006.
- Guix P, Parera M, Matas M, Castro JA, Ramon MM, Obrador A, Picornell A. Clinical implications of H63D homozygosity. *European Iron Club 2006* (Barcelona, Spain). 28-29 setembre 2006.
- Matas M, Picornell A, Cifuentes C, Payeras A, Bassa A, Homar F, López-Labrador FX, Moya A, Ramon MM, Castro JA. Relating the liver damage with hepatitis C virus polymorphism in HIV-1-coinfected patients. *12th Congress of the European Society for Evolutionary Biology 2010* (Turin, Italy).
- Castro JA, Matas M, Cifuentes C, Payeras A, Bassa A, Homar F, López-Labrador FX, Moya A, Picornell A, Ramon MM. The application of the Generalized Linear Model (GLM) to detect the main variables involved in liver damage caused by HCV/HIV coinfection. *13th Congress of the European Society for Evolutionary Biology 2011* (Tübingen, Germany).

