Journal of Viral Hepatitis, 2006, **13,** 104–115 doi:10.1111/j.1365-2893.2005.00670.x

Genetic similarity of hepatitis C virus and fibrosis progression in chronic and recurrent infection after liver transplantation

F.-X. López-Labrador,[1,2]* M. A. Bracho,[2]* M. Berenguer,[3] M. Coscollà,[2] J. M. Rayón,[4] M. Prieto,[3] D. Carrasco,[3] M. D. Gómez,[1] A. Moya[2] and F. González-Candelas[2] [1]*Microbiology/Exp. Immunology, Research Centre, Hospital Universitari La Fe;* [2]*Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València;* [3]*Hepatogastroenterology; and* [4]*Pathology Service, Hospital Universitari La Fe, Valencia, Spain*

Received February 2004; accepted for publication December 2004

SUMMARY. The effect of hepatitis C virus (HCV) genetic heterogeneity on clinical features of post-transplantation hepatitis C is controversial. Different regions of the HCV genome have been associated with apoptosis, fibrosis, and other pathways leading to liver damage in chronic HCV infection. Besides, differences in immunodominant regions, such as NS3, may influence HCV-specific immune responses and disease outcome. In the liver transplant setting, a recent study has reported a positive association between HCV-1b Core region genetic relatedness 5-year post-transplantation and histological severity of recurrent hepatitis C. We have compared nucleotide sequences of HCV Core, NS3 and NS5b regions in HCV-1b-infected patients 3 years post-transplantation ($n = 22$). A cohort of nontransplanted patients ($n = 22$) was used as control of natural chronic HCV-1b infection. Histological evaluation was used to define the rate of fibrosis progression. Molecular variance analysis did not show significant differences in HCV sequences between transplanted and nontransplanted patients, or between those with fast or slow fibrosis progression. The same results were obtained when analysing phylogenetic trees for Core, NS3 and NS5b regions. A more appropriate clustering method (using minimum spanning networks) revealed a significant positive relationship between HCV genetic similarity in Core ($r = 0.550$, $P < 0.01$) and NS5b regions ($r = 0.847$, $P < 0.01$) and the yearly rate of fibrosis progression in nontransplanted patients which, in contrast, was not observed in transplanted patients. Our results indicate that some strains of HCV-1b might be more pathogenic in the natural course of chronic infection by this virus subtype. In the liver transplant setting, when the immune response is severely compromised, other mechanisms are probably more important in determining hepatitis C progression.

Keywords: cirrhosis, disease progression, graft survival, minimum spanning network, phylogenetic tree, viral heterogeneity.

INTRODUCTION

Hepatitis C virus (HCV) infection is the major cause of chronic hepatitis, cirrhosis and end-stage liver disease leading to liver transplantation (OLT) world-wide. HCV contains a positive ssRNA of approximately 9600 nucleotides encoding a polyprotein of 3000–3010 amino acids. The HCV genome is characterized by a high replication error rate which leads to extensive genetic heterogeneity reflected in a number of distinct genotypes and subtypes [1], and in the so-called *quasispecies* distribution of the viral genome [2,3]. The degree of this variability depends on the genomic region analysed: from highly conserved regions (such as the 5′ end noncoding region – NCR-), regions with moderate variability (including the Core and nonstructural – NS- regions), and regions with a high degree of heterogeneity (such as the hypervariable region 1 – HVR-1 – in the envelope gene E2) [4].

Re-infection of the graft by HCV following OLT leads to histologically proven chronic hepatitis in the vast majority of patients [5]. The outcome of this infection is accelerated (with a faster progression of liver fibrosis) compared to that observed in nontransplanted patients [6,7]. Disease progression is however rather heterogeneous and, while some patients develop cirrhosis within 1 year of transplantation, others remain with stable histology for prolonged periods of time [8]. Prognosis in these patients is highly related to the evolution of HCV [9,10]. Because of the increasing need for

*These authors contributed equally to this work.

Abbreviations: GTR, Generalized Time Reversible; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; IVDU, intravenous drug use.

Correspondence: Dr Fernando González-Candelas, Institut Cavanilles de Biodiversitat i Biologia Evolutiva and Departament de Genètica, Universitat de València, Apartado Oficial 22085, 46071-València, Spain. E-mail: fernando.gonzalez@uv.es

organ donors the identification of factors associated with a fast histological progression post-OLT may have significant implications in liver transplant programmes. Degree of immunosupression, age of the donor, levels of viraemia prior to OLT, and HCV genotype are some of the variables which have been extensively evaluated in previous studies as potential determinants of outcome [5]. In contrast, few studies have focused on the genetic variability of HCV in liver transplant recipients with recurrent hepatitis C.

In the immunocompetent population, results on the association between the genetic heterogeneity of HCV and the outcome of hepatitis C are controversial [11–16]. In the OLT setting there are no conclusive data and the few published studies based on the analysis of the HVR-1 region in a small number of patients show discrepant results [17–20]. Emerging evidence shows that other HCV genomic regions may be implicated in the liver damage. In particular, HCV Core and NS3 proteins have been frequently associated with pathogenesis. Both Core and NS3 (and also NS4B and NS5A) can transform various cell lines, with or without the cooperation of oncogenes, and therefore have been linked to hepatocellular carcinoma (HCC) [21–25]. HCV Core protein has a modulatory role in programmed cell death under certain conditions, and Core genetic variability has been associated with differential NF-κB interaction [26,27]. Interestingly, a recent study analysed the genetic heterogeneity of the HCV Core region in cohorts of French and Italian liver transplant recipients 5 years post-OLT, and proposed that the fibrosis scores at this time point are related to the phylogenetic grouping of HCV Core sequences in patients infected by HCV subtype 1b [28].

Because liver damage in chronic hepatitis C is thought to be mainly caused by the host immune response against infected hepatocytes, the effects of several HCV proteins on the immune system has also been investigated. In this respect, it has been shown that HCV Core and NS3, but not E2 proteins, activate monocytes and myeloid dendritic cells [29]. The NS3 protein seems particularly immunogenic to the T-cell response, with relevant $CD4^+$ and $CD8^+$ T-cell epitopes among those mapped in the HCV polyprotein [30]. In particular, the $CD8^+$ T-cell epitopes within the NS3 protein are commonly recognized by T-cells in peripheral blood and liver of chronically infected patients [31,32]. Whereas T-cell reactivity against NS3 is associated with HCV clearance after acute infection, lack of reactivity, or an inappropriate response to this protein, is associated with the development of chronic hepatitis and fibrosis [30]. Furthermore, an association between outcome of HCV infection and T-cell escape amino acid variants in this genomic region has been reported in the chimpanzee model [33,34]. At the nucleotide level, recent studies suggest that some HCV-1b strains with particular NS3 sequences might be more frequently associated with fast fibrosis progression after OLT [35], or with the development of HCC in nontransplanted patients [36].

In the present study, we hypothesized that infection by phylogenetically different strains within HCV subtype 1b might explain the difference in outcome of hepatitis C post-OLT. Nucleotide sequences from the HCV Core, NS3 and NS5b regions were determined in serum samples from two cohorts of patients infected with HCV subtype 1b (22 liver transplant recipients, and 22 untreated, immunocompetent nontransplanted control patients) with well-characterized infection dates and yearly fibrosis progression rates based on liver biopsy examination. In both types of patients, we explored the genetic relationships between HCV isolates from patients with slow or fast fibrosis progression, defined by the rate of fibrosis progression from the date of infection.

METHODS

Patients

Two different populations of patients with chronic hepatitis C infected with HCV subtype 1b were enrolled in this study. The first group of patients (transplant group, $n = 22$) were selected among patients who had undergone OLT at our institution between January 1996 and January 2000 because of end-stage chronic hepatitis C. Inclusion criteria included HCV subtype 1b infection, positive HCV-RNA before and after liver transplantation, yearly biopsies available, lack of rejection episodes and no history of alcohol abuse. All of them had received standard immunosupressive regimes based on calcineurin inhibitors (azathioprine/tracolimus/prednisone $n = 2$; azathioprine/cyclosporine/prednisone $n = 4$; cyclosporine/prednisone $n = 2$; cyclosporine/mycophenolate/prednisone $n = 1$; tracolimus/prednisone $n = 13$). Only samples preantiviral therapy post-transplantation were used for the study. A second group of patients (nontransplant control group, $n = 22$) were selected on the basis of HCV subtype 1b infection from patients with chronic infection referred to our institution between 1999 and 2001 for antiviral treatment [37]. No patient had received antiviral treatment at the time of inclusion. Patients from both study populations had proven (serological, biochemical and histological) chronic hepatitis C, positive HCV-RNA by nested PCR; and negative hepatitis B surface antigen by MEIA (AXSYM HBsAg 2.0; Abbot, Weisbaden, Germany). Research protocols conformed to the ethical guidelines of the 1975 Declaration of Helsinki, and were approved by the Ethics Committee of our institution.

Rate of fibrosis progression

The rate of yearly fibrosis progression was calculated for each patient as the ratio between the stage of fibrosis observed on the last available liver biopsy and the number of years elapsed since the date of infection. Liver biopsies were performed routinely every year after transplantation. In

nontransplant patients, at least one liver biopsy was available in each patient, generally performed to evaluate therapy. All biopsy specimens were reviewed by a single pathologist in a blinded fashion, and were scored evaluating both the stage of fibrosis and the degree of necroinflammatory activity, as previously reported [8]. The grade was determined by combining the hepatic activity index scores for periportal necrosis, lobular degeneration and necrosis and portal inflammation, and was defined as follows: 1 to 2, minimal; 3 to 6, mild; 7 to 10, moderate; 11 to 14, severe. The different stages of fibrosis were defined as: 0, none; 1, fibrous portal expansion; 3, bridging fibrosis and 4, cirrhosis.

In liver transplant recipients, the date of transplantation was considered as the date of infection [38]. As previously shown [37,39], and in order to determine the date and route of infection in nontransplant patients, a careful interview and complete questionnaire regarding potential exposures was conducted. Risk factors for viral acquisition were classified into five major groups: (i) intravenous drug use (IVDU), (ii) transfusion, (iii) other known parenteral exposure (health care workers, tattoos, acupuncture, major surgery), (iv) family/sexual (history of HCV seropositive family member, history of multiple sexual partners) and (v) sporadic, when no risk factor was identified. When more than one risk factor was identified, the most likely one in the following scale was used for the analysis: (i) blood transfusion and IVDU, (ii) other known parenteral exposures, (iii) family/sexual. When more than one risk factor from the same category was present, the earliest potential exposure was used for the analysis. The date of presumed infection or the date of transplantation was used as time zero (fibrosis stage 0), and the fibrosis stage found in the last biopsy was divided by the time elapsed between the date of the biopsy and the presumed date of infection. Based on previous studies [37,39,40], patients were then classified as having slow or fast fibrosis progression (<0.1 or ≥ 0.1 units of fibrosis per year for nontransplant patients; <0.49 or ≥ 0.49 for liver transplant recipients, respectively).

Serum samples and virological tests

Samples were analysed at year 3 post-transplantation (transplant patients), or at the date of the first liver biopsy evaluation (nontransplant patients). For virological studies, blood samples were processed not later than 4 h after drawing, aliquoted, and frozen immediately at −70 °C. Blood samples from healthy anti-HCV seronegative individuals were used as negative controls. HCV-RNA was detected by means of nested reverse-transcription polymerase chain reaction (RT-PCR) of the 5′NCR as described [15]. The sensitivity of our nested PCR assay (10 HCV-RNA IU/mL) was estimated by testing serially 10-fold diluted samples in parallel with the Amplicor Monitor v2.0 assay (Roche Diagnostics, Barcelona, Spain). HCV-RNA quantitation was performed with the Amplicor HCV Monitor assay v2.0. All samples were first tested at either 1:10 or 1:100 dilution to prevent underestimation of values out of the dynamic range of the assay (600–500 000 IU/mL) [41]. HCV-RNA values were then corrected with the appropriate dilution factor in each case, and negative diluted samples were re-tested undiluted. HCV genotyping was determined by RFLP analysis of the 5′NCR as described previously [15,42], and confirmed on the basis of the NS5b sequence.

Amplification of the HCV Core, NS3 and NS5b regions

Briefly, total RNA was extracted from 140 μL of serum or plasma specimens with the QiaAmp HCV-RNA column kit (Qiagen GmbH, Hilden, Germany) following the instructions from the manufacturer. Reverse transcription was performed on a 20 μL volume containing 5 μL of eluted RNA, 4 μL of 5x RT buffer, 250 μM of each deoxynucleotide, 0.5 μg of random hexamers, 100 U of MMLV reverse transcriptase (Promega, Madison, WI, USA), and 20 U of RNasin Ribonuclease Inhibitor (Promega). The reaction was incubated at 42 °C, 1 min; 20 °C, 5 min; 25 °C, 5 min; 30 °C, 5 min; 35 °C 5 min; 37 °C, 30 min and then followed by 2 min at 95 °C. PCR was then carried out in a 50 μL volume containing 5 μL of the reverse transcription product, 200 μM of each dNTP, 0.4 μM of each primer and 0.5 U of *Pfu* DNA polymerase (Promega). Set primers for Core region (amplified product 387 bp) were Core-A1g (5′-GGGAGGTCTCGTAGACCGTGCACCATG-3′, 306–332) and Core-A1a (5′-GAGMGGKATRTACCCCATGAGRTCGGC-3′, 746–720) [primer positions refer to HCV isolate HCV-K1-R2 [43]; GeneBank accession number D50481]. PCR was performed in a thermal cycler (ABI 9700; Applied Biosystems, Foster City, CA, USA) according to the following profile: 94 °C, 1 min; 20 cycles at 94 °C, 30 s; 45 °C, 30 s; 72 °C, 30 s; then 20 cycles at 94 °C, 30 s; 60 °C, 30 s; 72 °C, 30 s and a final extension at 72 °C for 7 min. Set primers for the NS3 region (amplified product 517 bp) were NS3-5 (5′-ACGTACTCCACCTACGGCAA-3′; 4228–4248) and NS3-6 (5′-AAGGTAGGGTCAAGGCTGAA-3′; 4745–4765). PCR was performed according to the following profile: initial denaturation at 94 °C, 1 min; 3 cycles at 94 °C, 30 s; 55 °C, 30 s; 72 °C, 30 s, then 32 cycles at 94 °C, 30 s; 42 °C, 30 s; 72 °C, 30 s, and a final extension at 72 °C for 7 min, as previously described [35]. Set primers for NS5b region (amplified product 337 bp) were NS5b-A1g (5′-TATGATACYCGCTGYTTYGACTC-3′, 8241–8263) and NS5b-A1a (5′-GTACCTRGTCATAGCCTCCGTGAA-3′,8624–8601). PCR was performed according to the following profile: initial denaturation at 94 °C, 1 min; 5 cycles at 94 °C, 30 s; 55 °C, 30 s; 72 °C, 30 s, then 35 cycles at 94 °C, 30 s; 52 °C, 30 s; 72 °C 30 s, and a final extension at 72 °C for 7 min. In all three HCV amplified regions a single product was observed after electrophoresis on 1.4% agarose gels stained with ethidium bromide.

DNA sequencing

The PCR products were purified by using High Pure PCR Product Purification Kit (Roche Diagnostics GmbH, Mannheim, Germany). Purified DNA was gel-quantified, and 3–10 pg were used for direct sequencing with amplification primers and the ABI PRISM BigDye Terminator v3.0 Ready Reaction Cycle Sequencing KIT (Applied Biosystems) on an ABI 3700 Sequence Analyzer. Sequences were verified and both strands were assembled using the Staden package [44]. Sequences obtained have been deposited in GenBank with Accession Nos. AY898811 to AY898940.

Sequence comparison and phylogenetic analysis

Alignments of the nucleotide sequences were obtained with the ClustalW program [45]. Phylogenetic relationships among sequences were established using two different approaches. First, a phylogenetic tree was constructed using the neighbour-joining algorithm [46] based on the best evolutionary model for nucleotide substitution for each genomic region. Evolutionary models were evaluated using Modeltest [47] and the one maximizing the AIC value [48] was selected. Bootstrap resampling (2000 replicates) was used to measure the statistical support for each node in the phylogenetic trees. These analyses were done with PAUP* v4b10 [49]. Alternatively, phylogenetic relationships were established using minimum spanning networks [50] with the absolute number of differences between sequences using Arlequin 2.000 [51].

Statistical analysis

Comparisons of variables between patient populations and groups were performed by two-sided *t*-tests (95% confidence interval) or chi-square tests. Correlation analyses were performed by conventional methods. Statistical calculations were performed with the SPSS v11.0 package. Comparisons between genetic distances of HCV-isolates were made using the analysis of molecular variance [52] as implemented in Arlequin 2.000 [51]. This analysis allows one to compare the distribution of genetic variation among subsets of sequences defined *a priori* and to establish whether these groups are significantly different from each other.

RESULTS

Patient characteristics

The main characteristics of the patient population are summarized in Table 1. Transplanted patients were significantly older, had shorter period postinfection, and higher fibrosis progression rate than nontransplanted patients. On the basis of yearly rates of fibrosis progression, 12 from the 22 transplanted patients (54.5%) were classified as slow progressors, and the remaining 10 as fast progressors. Fast progressors had significantly higher mean activity grade and mean fibrosis scores in their last available biopsy than slow progressors. Thirteen from the 22 nontransplanted patients (59.1%) were classified as slow progressors, and the remaining nine as fast progressors. No other significant differences were found between slow or fast progressors in both patient populations.

HCV-RNA and viral load

Only one sample tested negative with the Amplicor Monitor v2.0 test, but positive with our qualitative 'in-house' RT-PCR assay, and an arbitrarily assigned value of 50 HCV-RNA IU/mL was used for data analysis. Mean viral load was slightly higher in transplanted patients than in the nontransplanted group, but this difference lacked statistical significance (Table 1).

HCV nucleotide sequence analysis

Hepatitis C virus nucleotide sequences did not show any stop codons, insertions or deletions. Our first analysis of the relationship between liver fibrosis progression and genetic relatedness of HCV did not consider the phylogeny of the sequences involved. We first divided the patients in different groups according to the rate of fibrosis progression and transplant status. Analysis of molecular variance was used to compare the HCV nucleotide sequences in these groups (Table 2). None of the comparisons between groups, fast *vs* slow progressors and transplanted *vs* nontransplanted, for the three HCV genomic regions was significant, thus indicating that viral isolates from any of these groups were not significantly different from the others in the comparison. Similar results were obtained when groups were defined according to fibrosis score (mild: fibrosis 0–1; severe: fibrosis 3–4) and transplant status (data not shown). No particular pattern of amino acid sequences in relevant NS3 T-cell epitopes (LIFCHSKKK$_{1391}$ CHSKKKCDE$_{1395}$, ELAAKLVAL$_{1402}$, KLVALGINAV$_{1406}$ and ATDALMTGF$_{1435}$) was associated to fast or slow fibrosis progression (data not shown). We then proceeded to evaluate the phylogenetic relationships among the sequences derived from each individual patient.

HCV phylogenetic analysis

For the HCV Core region, the analysis of 56 alternative evolutionary models of nucleotide substitution resulted in the selection of a Generalized Time Reversible (GTR) model [53] with a proportion of 68.81% invariant sites and a gamma distribution with shape parameter 0.4623 accounting for the heterogeneity rate among nucleotide sites. The same analysis of the NS3 region resulted in the selection of a transversion model [54], with 56.55% of invariant sites and a gamma distribution with shape parameter 2.1972. For the NS5b region the analysis also resulted in the selection of a

Table 1 Main characteristics of transplant and nontransplant patients

	Age ($\bar{x}$ ± SD)	Gender (male,%)	Years p.i. [a]($\bar{x}$ ± SD)	Activity grade ($\bar{x}$ ± SD)	Fibrosis stage ($\bar{x}$ ± SD)	Fibrosis progression ($\bar{x}$ ± SD)	ALT ($\bar{x}$ ± SD)	Viral load ($\bar{x}$ $\log_{10}$ IU/mL ± SD)
Transplant ($n = 22$)	58 ± 8.8[b]	14 (63.6)	2.50 ± 1.33[c]	7.55 ± 4.11	1.95 ± 1.94	1.56 ± 1.79[h]	91.32 ± 63.85	6.01 ± 0.93
Slow progression ($n = 12$)	61 ± 8.0	6 (50)	2.93 ± 1.07	4.42 ± 2.19[f]	0.25 ± 0.45[g]	0.13 ± 0.23[i]	78.33 ± 44.52	6.32 ± 0.59[k]
Fast progression ($n = 10$)	54 ± 8.3	8 (80)	1.97 ± 1.47	11.30 ± 2.21	4.00 ± 0	3.27 ± 1.19	106.90 ± 81.21	5.64 ± 1.14
Nontransplant ($n = 22$)	47 ± 12.3	13 (59.1)	21.64 ± 13.20[d]	7.00 ± 2.81	1.82 ± 1.05	0.18 ± 0.25	77.68 ± 60.10	5.66 ± 1.12
Slow progression ($n = 13$)	49 ± 8.8	9 (69.2)	28.57 ± 11.29[e]	7.46 ± 3.26	1.54 ± 1.05	0.05 ± 0.03[j]	66.62 ± 50.51	5.81 ± 1.32
Fast progression ($n = 9$)	43 ± 15.9	4 (44.4)	11.62 ± 8.64	6.33 ± 2.00	2.22 ± 0.97	0.35 ± 0.33	93.67 ± 71.89	5.45 ± 0.78

NA, nonapplicable; M, male; F, female.
[a]Years postinfection (p.i.) correspond to years post-transplantation in transplanted patients; [b]$P < 0.01$ *vs* Nontransplant (two-tailed *t*-test); [c]$P < 0.001$ *vs* nontransplant (two-tailed *t*-test); [d]$P < 0.01$ *vs* fast progression (two-tailed *t*-test); [e]$P \leq = 0.01$ *vs* fast progression (two-tailed *t*-test); [f]$P < 0.001$ *vs* fast progression (Mann–Whitney *U*-test); [g]$P < 0.001$ *vs* fast progression (Mann–Whitney *U*-test); [h]$P < 0.01$ *vs* nontransplant (two-tailed *t*-test); [i]$P < 0.001$ *vs* fast progression (two-tailed *t*-test); [j]$P < 0.05$ *vs* fast progression (two-tailed *t*-test); [k]$P < 0.10$ *vs* fast progression.

Table 2 Summary of the molecular variance analyses (AMOVA) for the Core and NS5b regions of HCV sequences derived from patients included in this study

	Core					NS3					NS5b				
Source of variation	d.f.	Sum of squares	Variance components	% total variation	*P*-value	d.f.	Sum of squares	Variance components	% total variation	*P*-value	d.f.	Sum of squares	Variance components	% total variation	*P*-value
Among groups (s-f)	1	4.485	0.001	0.01	ns	1	8.780	−0.281	−2.63	ns	1	6.893	−0.257	−2.62	ns
Among populations within groups (TOH-IC)	2	8.941	−0.043	−0.88	ns	2	29.296	0.389	3.64	ns	2	4.961	0.244	2.49	ns
Within populations	40	197.846	4.94615	100.88		38	401.686	10.571	98.99		40	392.260	9.807	100.14	
Total	43	211.273	4.902			41	439.762	10.679			43	424.114	9.790		
Among groups (TOH-IC)	1	5.045	0.040	0.82	ns	1	15.389	0.274	2.53	ns	1	13.114	0.171	1.72	ns
Among populations within groups (s-f)	2	8.381	−0.070	−1.43	ns	2	19.398	−0.093	−0.8	ns	2	18.740	−0.041	−0.41	ns
Within populations	40	197.846	4.946	100.61		38	404.975	10.657	98.34		40	392.260	9.807	100.14	
Total	43	211.273	4.916			41	439.762	10.837			43	424.114	9.937		

s-f, slow *vs* fast progressors; TOH-IC, transplanted *vs* nontransplanted; ns, nonsignificant.

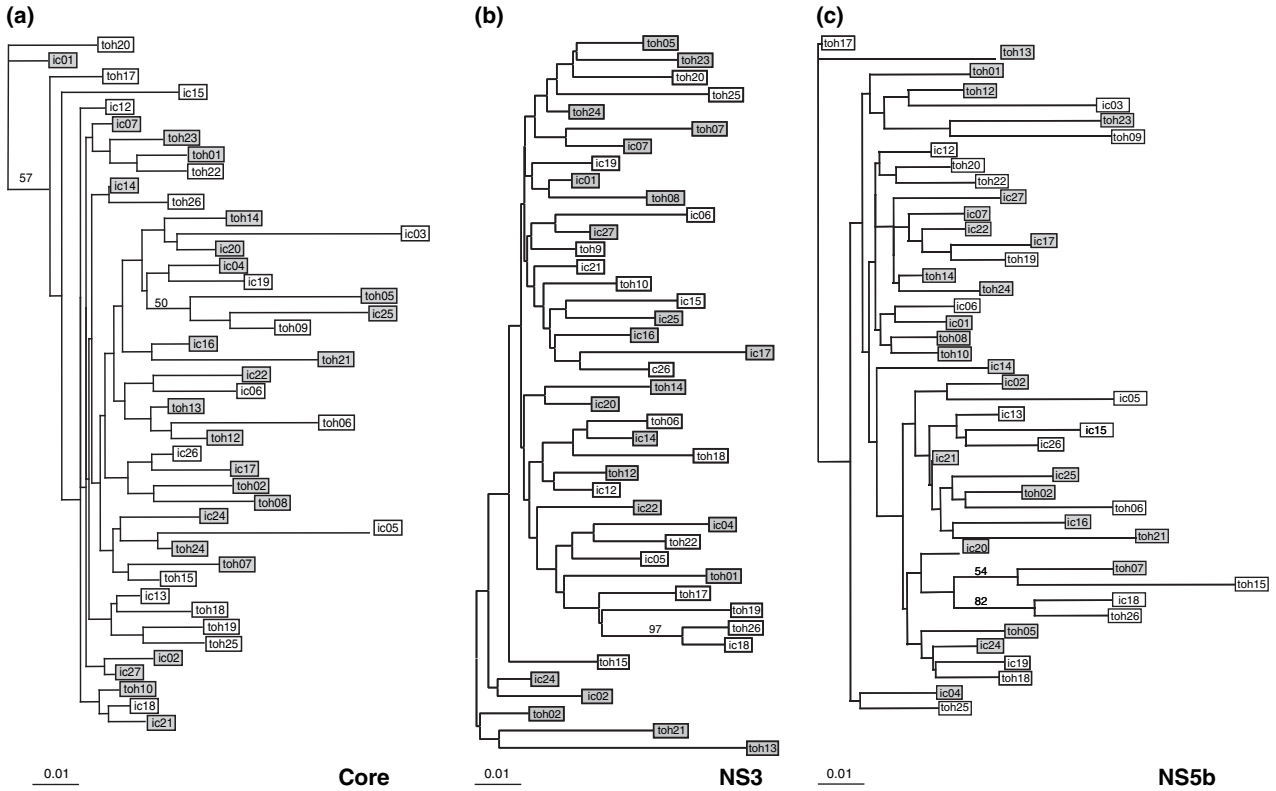

Fig. 1 Phylogenetic trees of (a) Core, (b) NS3, and (c) NS5b regions obtained with the neighbour-joining algorithm after the most appropriate evolutionary model was determined by maximum likelihood analysis. Bootstrap values > 50% are shown in the corresponding nodes. White boxes indicate sequences from fast progressors; shaded boxes correspond to slow progressors. Scale bars correspond to nucleotide substitutions per site. toh: transplanted; ic: nontransplanted (immunocompetent).

transversion model with 59.09% of invariant sites and a gamma distribution with shape parameter 0.6391. The phylogenetic trees obtained for the Core, NS3 and NS5b regions are shown in Fig. 1. Two remarkable features can be observed in these trees. First, most nodes have very low statistical support since only two nodes in Core and NS5b trees, and one node in the NS3 tree show bootstrap support values higher than 50%. Besides, the supported nodes are different for the three genomic regions analysed. This lack of congruence between the trees extends to the remaining nodes and, as a consequence, prevented us from considering any of them (not even a combined analysis of the three regions), as reliable representations of the phylogenetic relationships among the sequences. For the three genomic regions, the sequences of this study were dispersed in phylogenetic trees among a panel of control sequences isolated from unrelated Spanish patients (data not shown), thus indicating that they are not too closely related nor derived from a common source (i.e. they do not belong to an outbreak).

Minimum spanning networks

The lack of phylogenetic resolution prevented us from analysing the relationship between fibrosis progression and sequence relatedness defined from a phylogenetic tree. In order to obtain a better working definition of relatedness, we used an alternative clustering method by constructing a minimum spanning network from the matrix of the absolute number of differences between each pair of sequences for each genomic region. The resulting networks are shown in Fig. 2. The three genomic regions present a similar pattern with many alternative connections, 22 for the Core region, 14 for NS3, and 19 for NS5b. These alternative connections explain the low support for the phylogenetic trees, since the algorithms used in tree construction try to obtain one single, strictly bifurcating representation of what actually is a rather tangled network. However, using this approach, it is possible to define all pairs of most similar sequences and then proceed with the analysis.

We first checked whether viral isolates from fast or slow progressors, regardless of their transplant status, were more closely related than expected by chance. Next, we computed the number of connections between isolates from slow–slow, fast–fast and slow–fast progressors and compared them with the expected numbers under the null distribution of random associations. For the three HCV genomic regions, the chi-square tests were not statistically significant ($\chi^2 =$ 0.1198, 1.6916 and 2.8546, with 1 d.f. for Core, NS3 and

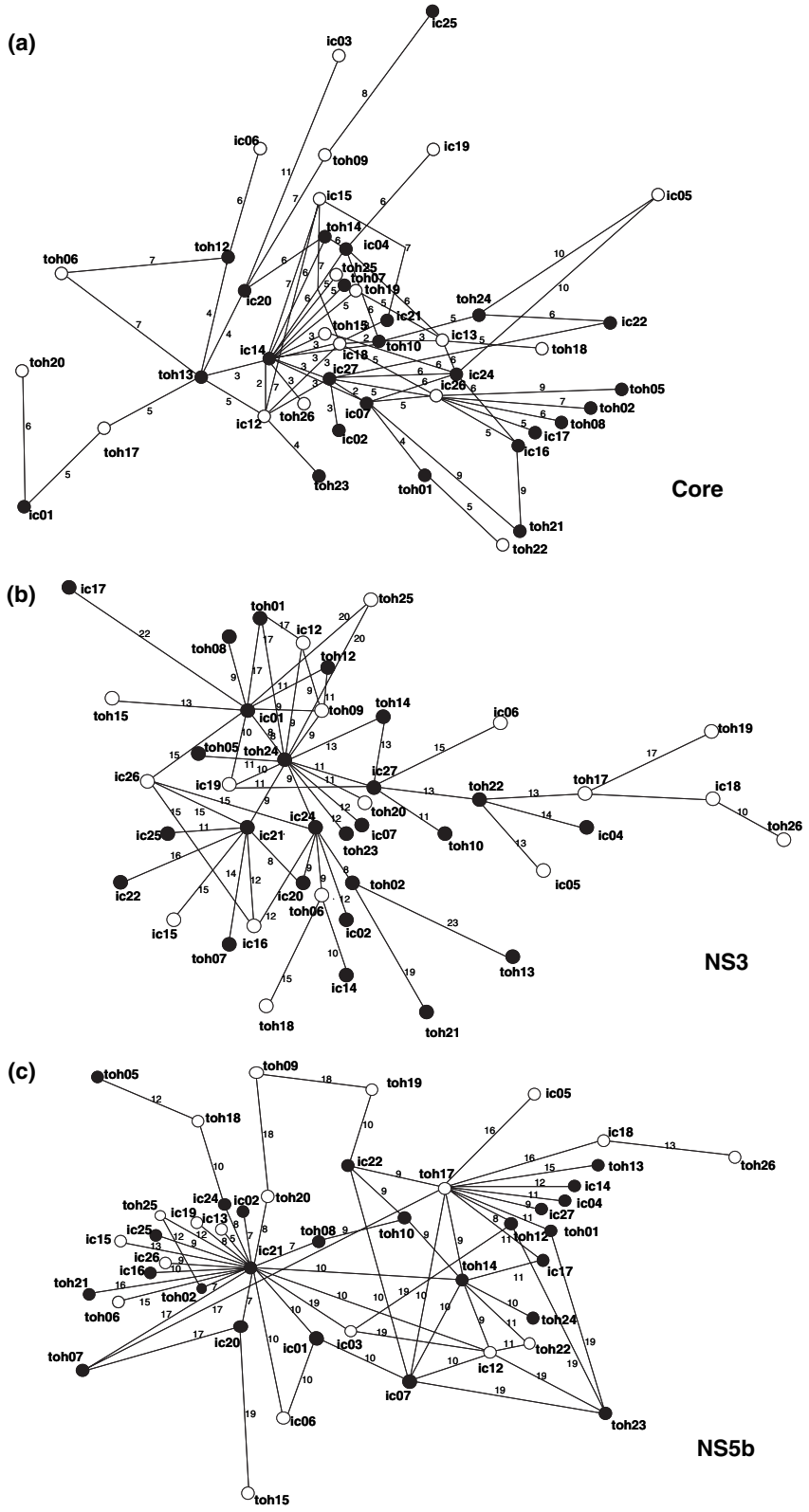

Fig. 2 Minimum spanning network for the (a) hepatitis C virus (HCV) Core, (b) NS3, and (c) NS5b regions constructed from the matrix of the absolute number of differences between each pair of sequences. Figures denote the number of nucleotide differences between sequences at both ends of each segment. Open circles: HCV sequences from fast progressors; solid circles: HCV sequences from slow progressors; toh: transplanted; ic: nontransplanted (immunocompetent).

NS5b, respectively), thus suggesting that there is no evidence of a closer (or farther) relatedness between similar or different types of progressors. The same result was obtained when the comparisons were made on the basis of fibrosis scores ($X^2 = 0.3826$, 1.8233 and 2.6145 for Core, NS3 and NS5b regions, respectively).

A slightly different question is whether similar or different types of progressors are positively or negatively correlated with regards to the number of differences in HCV sequences (nucleotide substitutions). We checked this possibility in two ways. First, we used the same procedure described above but this time computing the number of differences between adjacent pairs of sequences in each network. We next computed the total number of differences for each class of pairs of progressors (slow–slow, fast–fast and slow–fast) and compared them with the expected numbers under the null hypothesis of identical average differences for the three classes. For the Core region, the chi-square test was not significant ($\chi^2 = 1.707$, 1 d.f., $P > 0.10$), but for the NS3 and NS5b regions the test showed a significant value ($\chi^2 = 9.9553$, $P < 0.01$; and $\chi^2 = 4.207$, $P = 0.04$, respectively, with 1 d.f.) due to a slightly larger number of differences between fast progressors than under the null hypothesis (the average observed distance was 13.90 while the expected value was 11.98 for NS5b, and 13.4 and 13.09, respectively, for NS3). However, it should be taken into account that this is the less common class of pairs, with only 10 cases from a total of 62 pairs. The remaining two classes presented a different pattern, with fewer differences than expected: the average distances between close pairs of slow progressors were 11.30 for NS5b and 12.13 for NS3 and between fast–slow pairs the corresponding observed values were 11.86 for NS5b and 13.41 for NS3, both values being very similar to the expected, 11.98 and 13.74 respectively, differences on average. Similar results and conclusions were obtained when the fibrosis score from each individual rather than its yearly rate of progression was used to classify the patients.

A second analysis involved the correlation between the number of nucleotide differences between neighbour sequences and the absolute difference in the yearly rate of fibrosis progression of the corresponding patients. Again, only pairs of sequences connected in the minimum spanning networks were considered. Overall, for the three genomic regions the correlations were very low or nonsignificant, with Pearson correlation coefficients of 0.078 for the Core region ($n = 65$), 0.1111 for NS3 ($n = 55$) and 0.00075 for the NS5b region ($n = 62$). However, when the comparisons took into account the transplant status of each patient, a different picture emerged. For nontransplanted individuals, a positive significant correlation was found between the absolute value of the difference in the rate of progression and the number of nucleotide substitutions for both the Core and NS5b HCV genomic regions (Table 3). These correlations were not significant for transplanted individuals. A similar pattern was obtained when patients with low and high fibrosis scores were compared (data not shown). These data suggest that in nontransplanted patients there is a positive correlation between HCV genetic similarity and similarity in rate of fibrosis progression, whereas this correlation is not observed in transplanted patients.

Table 3 Correlation between genetic similarity and rate of fibrosis progression for transplant and nontransplant patients. For each genomic region, only those pairs of individuals with matching transplant status and connected in the corresponding minimum spanning network (Fig. 2) were considered. The absolute value of the difference between rates of fibrosis progression was compared with the minimum number of nucleotide substitutions in the respective sequences

	Genomic region		
Patient group	Core	NS3	Ns5b
Nontransplanted	$r = 0.550$	$r = 0.218$	$r = 0.847$
	$n = 27$	$n = 16$	$n = 18$
	$P < 0.01$	$P > 0.10$	$P < 0.01$
Transplanted	$r = 0.094$	$r = 0.1489$	$r = -0.332$
	$n = 6$	$n = 8$	$n = 15$
	$P > 0.10$	$P > 0.10$	$P > 0.10$

r, Spearman's correlation coefficient; n, number of pairs compared; P, probability for $r = 0$.

DISCUSSION

The natural history of hepatitis C is variable, both in transplanted and in nontransplanted patients. Studies evaluating the effects of HCV genetic heterogeneity on outcome do not clarify this issue. Our study was aimed at establishing a relationship between genetic similarity of the HCV-1b infecting strains and recurrent HCV disease progression. In light of recent data [28,35], our hypothesis was that some HCV-1b strains are associated with a higher yearly rate of fibrosis progression post-transplantation. A group of untreated, nontransplant patients was included as control for the natural course of chronic HCV infection. Our results may be summarized as follows: (i) molecular variance and conventional phylogenetic analysis of the HCV Core, NS3 and NS5b regions yielded insufficient resolution for differentiating subtype 1b strains infecting fast and slow progressors; (ii) a more complete analysis (minimum spanning networks) allows pairwise comparison of HCV sequences only between relevant patient pairs (i.e. slow–fast). This approach reveals a significant relatedness of Core and NS5b regions with similarity in fibrosis progression in HCV-1b isolates from nontransplanted patients but not in isolates from transplanted patients.

Previous studies have found discrepant results regarding the effect of HCV heterogeneity on outcome of infection. Reasons for these discrepancies may be: (i) the analysis of different HCV genomic regions; (ii) definitions of both disease severity (disease severity or progression, survival, rate of recurrence) and HCV heterogeneity (*quasispecies* complexity, phylogenetic clustering) and (iii) heterogeneity of patient populations from different geographical areas infected by

different HCV subtypes with potentially different pathogenic effects [15]. Indeed, the majority of studies have analysed HCV evolution in the HVR-1 [17,19,20,55], although some are based in other genomic regions [28,56]. In addition, absolute fibrosis scores, which represent a static information, have been used as a measure of the progression of chronic hepatitis C, which is a dynamic process. Finally, definition of HCV heterogeneity also differs between studies. Most have evaluated HCV *quasispecies* diversity on outcome [17,19,20,56]. A different approach is to define whether the phylogenetic similarity between different HCV strains influences the severity of liver disease [28].

In our study, we did not pretend to analyse the HCV *quasispecies* diversity, but rather the genetic relatedness of HCV between strains infecting patients with slow or fast fibrosis progression. For this purpose, we used a more complete alternative genetic analysis than those typically performed, and we evaluated the yearly rate of fibrosis progression, a measure which considers liver fibrosis as a dynamic variable changing along with time. This approach avoids the difficulties which arise by using a categorical variable with arbitrary integer values, such as the fibrosis score, in linear correlation analyses with HCV genetic quantitative variables. Furthermore, instead of a simple phylogenetic tree construction, our genetic analysis explores HCV-1b genetic features in the Core, NS3 and NS5b regions in a step-wise manner, so that more complex genetic relationships between HCV-1b strains can be established depending on the results obtained in prior steps. In the first step, phylogenetic grouping of HCV-1b isolates (indicating common evolutionary origin) was not possible. Thus, only well-established significant genetic relationships found in the minimal spanning networks were considered. This approach avoids the confounding effect of considering only a few from many alternative relationships among sequences, which reflects the uncertainty in the phylogenetic reconstruction with highly variable viral sequences. Finally, only patients infected with HCV subtype 1b were included in our study, thus eliminating the HCV subtype as a confounding factor on disease progression, severity and outcome.

When all the sequences were analysed together, there was not a higher genetic variability of HCV in patients with different rates of fibrosis progression, neither in transplanted or in nontransplanted patients. Overall, HCV-1b strains from patients with slow (or fast) fibrosis progression are not more similar among themselves than to the remaining sequences. When analysing the results according to transplant status, a significant positive correlation was found between HCV-1b genetic similarity and the yearly rate of fibrosis progression in nontransplanted individuals, in both Core and NS5b regions, albeit not detected by the usual phylogenetic analysis methods. These results might be explained by infection with different HCV-1b strains with potentially different pathogenic effects, although another study [57] has found different degrees of liver damage (mild and severe) 20 years postinfection in a cohort infected in an outbreak with a single HCV-1b isolate. Unfortunately, disease progression was not determined in that study.

Surprisingly, we were unable to correlate HCV-1b genetic similarity and rate of fibrosis progression in transplant patients. These results are in contrast with the finding of a significant phylogenetic subgrouping of HCV genotype 1 strains with respect to the fibrosis scores [28]. The reasons for this discrepancy may be as follows. First, the year of transplantation was 1987–1995 in the study by Gigou *et al.*, whereas patients included in our study were transplanted in the 1996–2000 period. This difference may implicate differences in the use of stronger immunosupressing drugs, and therefore in fibrosis progression, between both cohorts [40]. Indeed, in our study 15 of 22 (68%) patients received tracolimus-based immunosupression, compared with 27 of 68 (39%) patients in the study by Gigou *et al.* Secondly, our study focused in subtype 1b infection only. Thirdly, we used the GTR model to calculate Core genetic distances after testing for the best evolutionary model. Finally, in the study by Gigou *et al.*, fibrosis was analysed at year 5 post-OLT, whereas we analysed at year 3 post-OLT. It is possible that some of our patients with low fibrosis scores (and/or slow fibrosis progression) develop more severe liver disease at year 5 post-OLT [58]. Further follow-up of our cohort is certainly warranted.

In a pilot study including a small number of patients, we previously showed some phylogenetic grouping, albeit with no significant clustering, in the NS3 region of HCV-1b strains from transplanted and nontransplanted HCV-infected patients with fast fibrosis progression [35]. Why we failed to associate HCV genetic similarity in NS3 with similarity of fibrosis progression in the current study (neither in transplanted nor in nontransplanted patients) can be due to the increased sample size. Furthermore, no particular pattern of amino acid sequences in relevant NS3 T-cell epitopes was associated to fast or slow fibrosis progression. In contrast, the correlation between HCV genetic similarity in Core and NS5b with similarity of fibrosis progression in nontransplanted patients presumably indicates a pathogenic role for this two HCV proteins.

Our data could be interpreted in two ways. First, that there are no HCV-1b strains with differential pathogenic capacity infecting our population of liver transplant recipients. Alternatively, the 'more pathogenic' features of particular HCV-1b strains (which might lead to increased fibrosis progression in nontransplant patients), may be less relevant in the liver transplant area. After OLT, a completely different immunological scenario emerges (i.e. organ-donor HLA-mismatch, depression of the cellular immune response by immunosupressive drugs and, therefore, less immune pressure). In this particular setting, other mechanisms, related either to the virus or to the host, may be more important in determining the outcome of recurrent

hepatitis C. Eventually, analysis of sequence diversity and complexity of HCV in patients from different geographical regions will most likely be necessary to test the hypothesis that some strains of HCV subtype 1b are associated with disease progression in transplanted or nontransplanted patients.

In our study, HCV RNA levels were found not to be significantly different between transplanted and immunocompetent individuals. This finding is at odds with reports from studies showing significant increase in viraemia levels following liver transplantation [59]. Reasons which may explain these discrepancies include (i) the use of a dilution step applied to all samples in our study, in order to obtain an HCV-RNA estimation within the linear range of the Amplicor Monitor Assay, (ii) the time-points at which samples are analysed. In our study, the analysis was performed 3-year post-transplantation (median), when HCV-RNA levels may be lower because of decreased immunosupression and reconstitution of the host cellular immune response and (iii) 45.5% of our transplanted patients had already progressed to a cirrhotic stage at the time of sampling, thus potentially with less hepatocytes supporting viral replication [28].

In summary, we found a significant positive correlation between the genetic similarity of HCV-1b strains (as determined by minimal spanning network analysis) and the yearly rate of fibrosis progression in immunocompetent patients. These results suggest that different HCV-1b strains with differential pathogenic potential may be implicated in the natural history of hepatitis C in our geographical area. In the liver transplant setting however, when the immune system is severely compromised and the virus replicates under different conditions, other factors are probably more relevant in determining the outcome of the disease.

DEDICATION

We wish to dedicate this work to our colleague and friend Dr Domingo Carrasco, who sadly passed away during the revision process of this paper. Dr Carrasco served as staff physician in the HepatoGastroenterology Service of our Institution from the very beginning of the Service, in the 1960s. He had always been an enthusiastic and outstanding person both at the personal and professional level. It was his enthusiasm and personal integrity which made him continue working in the face of a long-lasting disease.

ACKNOWLEDGEMENTS

We are indebted to José V. Sarrión (Hepatogastroenterology Service, Hospital Universitari La Fe) for patient evaluation and to M[a] José Giménez (Microbiology Service, Hospital Universitari La Fe) for serological determinations of hepatitis markers. Sequences were obtained in the Servicio de Secuenciación, SCSIE, Universitat de València. This work was supported by grants from the Fondo de Investigación Sanitaria, Instituto de Salud Carlos III Ministerio de Sanidad y Consumo (Projects PI01/1196, and C03/02 – Red Nacional de Investigación en Hepatología y Gastroenterología); Ministerio de Ciencia y Tecnología (Project BMC2001–3096); and Conselleria de Sanitat de la Generalitat Valenciana. F-X.L-L. holds a position from the Spanish Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Ministerio de Sanidad y Consumo (contract numbers 99/3101 and CP04100020).

REFERENCES

1 Simmonds P, Alberti A, Alter HJ *et al.* A proposed system for the nomenclature of hepatitis C viral genotypes. *Hepatology* 1994; 19(5): 1321–1324.
2 Domingo E, Escarmís C, Sevilla N *et al.* Basic concepts in RNA viruses evolution. *FASEB J* 1996; 10: 859–864.
3 Martell M, Esteban JI, Quer J *et al.* Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J Virol* 1992; 66(5): 3225–3229.
4 Forns X, Bukh J. The molecular biology of hepatitis C virus. Genotypes and quasispecies. *Clin Liver Dis* 1999; 3: 693–716.
5 Berenguer M, López-Labrador FX, Wright TL. Hepatitis C and liver transplantation. *J Hepatol* 2001; 35(5): 666–678.
6 Berenguer M, Prieto M, Rayón JM *et al.* Natural history of clinically compensated hepatitis C virus-related graft cirrhosis after liver transplantation. *Hepatology* 2000; 32(4 Pt 1): 852–858.
7 Poynard T, Bedossa P, Opolon P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. *Lancet* 1997; 349(9055): 825–832.
8 Berenguer M, Rayón JM, Prieto M *et al.* Are posttransplantation protocol liver biopsies useful in the long term? *Liver Transpl* 2001; 7(9): 790–796.
9 Berenguer M, Prieto M, San Juan F *et al.* Contribution of donor age to the recent decrease in patient survival among HCV-infected liver transplant recipients. *Hepatology* 2002; 36(1): 202–210.
10 Prieto M, Berenguer M, Rayón JM *et al.* High incidence of allograft cirrhosis in hepatitis C virus genotype 1b infection following transplantation: relationship with rejection episodes. *Hepatology* 1999; 29(1): 250–256.
11 Curran R, Jameson CL, Craggs JK *et al.* Evolutionary trends of the first hypervariable region of the hepatitis C virus E2 protein in individuals with differing liver disease severity. *J Gen Virol* 2002; 83(1): 11–23.
12 Honda M, Kaneko S, Sakai A, Unoura M, Murakami S, Kobayashi K. Degree of diversity of hepatitis C virus quasispecies and progression of liver disease. *Hepatology* 1994; 20(5): 1144–1151.
13 Koizumi K, Enomoto N, Kurosaki M *et al.* Diversity of quasispecies in various disease stages of chronic hepatitis C virus infection and its significance in interferon treatment. *Hepatology* 1995; 22(1): 30–35.
14 González-Peralta RP, Qian K, She JY *et al.* Clinical implications of viral quasispecies heterogeneity in chronic hepatitis C. *J Med Virol* 1996; 49(3): 242–247.

15 López-Labrador FX, Ampurdanés S, Forns X *et al.* Hepatitis C virus (HCV) genotypes in Spanish patients with HCV infection: relationship between HCV genotype 1b, cirrhosis and hepatocellular carcinoma. *J Hepatol* 1997; 27(6): 959–965.
16 Naito M, Hayashi N, Moribe T *et al.* Hepatitis C viral quasispecies in hepatitis C virus carriers with normal liver enzymes and patients with type C chronic liver disease. *Hepatology* 1995; 22(2): 407–412.
17 Gretch DR, Polyak SJ, Wilson JJ, Carithers RL Jr, Perkins JD, Corey L. Tracking hepatitis C virus quasispecies major and minor variants in symptomatic and asymptomatic liver transplant recipients. *J Virol* 1996; 70(11): 7622–7631.
18 Lawal Z, Petrik J, Wong VS, Alexander GJ, Allain JP. Hepatitis C virus genomic variability in untreated and immunosuppressed patients. *Virology* 1997; 228(1): 107–111.
19 Pessoa MG, Bzowej N, Berenguer M *et al.* Evolution of hepatitis C virus quasispecies in patients with severe cholestatic hepatitis after liver transplantation. *Hepatology* 1999; 30(6): 1513–1520.
20 Sánchez-Fueyo A, Giménez-Barcons M, Puig-Basagoiti F *et al.* Influence of the dynamics of the hypervariable region 1 of hepatitis C virus (HCV) on the histological severity of HCV recurrence after liver transplantation. *J Med Virol* 2001; 65(2): 266–275.
21 Ray RB, Lagging LM, Meyer K, Ray R. Hepatitis C virus core protein cooperates with ras and transforms primary rat embryo fibroblasts to tumorigenic phenotype. *J Virol* 1996; 70(7): 4438–4443.
22 Ray RB, Meyer K, Ray R. Hepatitis C virus core protein promotes immortalization of primary human hepatocytes. *Virology* 2000; 271(1): 197–204.
23 Sakamuro D, Furukawa T, Takegami T. Hepatitis C virus nonstructural protein NS3 transforms NIH 3T3 cells. *J Virol* 1995; 69(6): 3893–3896.
24 Park JS, Yang JM, Min MK. Hepatitis C virus nonstructural protein NS4B transforms NIH3T3 cells in cooperation with the Ha-ras oncogene. *Biochem Biophys Res Commun* 2000; 267(2): 581–587.
25 Ghosh AK, Steele R, Meyer K, Ray R, Ray RB. Hepatitis C virus NS5A protein modulates cell cycle regulatory genes and promotes cell growth. *J Gen Virol* 1999; 80(Pt 5): 1179–1183.
26 Ray RB, Ray R. Hepatitis C virus core protein: intriguing properties and functional relevance. *FEMS Microbiol Letters* 2001; 202(2): 149–156.
27 Ray RB, Steele R, Basu A *et al.* Distinct functional role of Hepatitis C virus core protein on NF-κB regulation is linked to genomic variation. *Virus Res* 2002; 87(1): 21–29.
28 Gigou M, Roque-Afonso AM, Falissard B, Penin F, Dussaix E, Féray C. Genetic clustering of hepatitis C virus strains and severity of recurrent hepatitis after liver transplantation. *J Virol* 2001; 75(23): 11292–11297.
29 Dolganiuc A, Kodys K, Kopasz A *et al.* Hepatitis C virus core and nonstructural protein 3 proteins induce pro- and anti-inflammatory cytokines and inhibit dendritic cell differentiation. *J Immunol* 2003; 170(11): 5615–5624.
30 Racanelli V, Rehermann B. Hepatitis C virus infection: when silence is deception. *Trends Immunol* 2003; 24(8): 456–464.
31 He XS, Rehermann B, López-Labrador FX *et al.* Quantitative analysis of hepatitis C virus-specific CD8(+) T cells in peripheral blood and liver using peptide-MHC tetramers. *Proc Natl Acad Sci USA* 1999; 96(10): 5692–5697.
32 López-Labrador FX, He XS, Berenguer M *et al.* The use of class-I HLA tetramers for the detection of hepatitis C virus NS3-specific CD8(+) T cells in patients with chronic infection. *J Immunol Methods* 2004; 287(1–2): 91–99.
33 Weiner A, Erickson AL, Kansopon J *et al.* Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escae variant. *Proc Natl Acad Sci USA* 1995; 92: 2755–2759.
34 Erickson AL, Kimura Y, Igarashi S *et al.* The outcome of hepatitis C virus infection is predicted by escape mutations in epitopes targeted by cytotoxic T lymphocytes. *Immunity* 2001; 15(6): 883–895.
35 López-Labrador FX, Berenguer M, Sempere A *et al.* Genetic variability of hepatitis C virus NS3 protein in human leukocyte antigen-A2 liver transplant recipients with recurrent hepatitis C. *Liver Transpl* 2004; 10(2): 217–227.
36 Ogata S, Huab Florese R, Nagano-Fujii M *et al.* Identification of hepatitis C virus (HCV) subtype 1b strains that are highly, or only weakly, associated with hepatocellular carcinoma on the basis of the secondary structure of an amino-terminal portion of the HCV NS3 protein. *J Clin Microbiol* 2003; 41(7): 2835–2841.
37 Ortiz V, Berenguer M, Rayon JM, Carrasco D, Berenguer J. Contribution of obesity to hepatitis C-related fibrosis progression. *Am J Gastroenterol* 2002; 97(9): 2408–2414.
38 García-Retortillo M, Forns X, Feliu A *et al.* Hepatitis C virus kinetics during and immediately after liver transplantation. *Hepatology* 2002; 35(3): 680–687.
39 Poynard T, Bedossa P, Opolon P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet* 1997; 349(9055): 825–832.
40 Berenguer M, Ferrell L, Watson J *et al.* HCV-related fibrosis progression following liver transplantation: increase in recent years. *J Hepatol* 2000; 32(4): 673–684.
41 Pawlotsky JM. Molecular diagnosis of viral hepatitis. *Gastroenterology* 2002; 122(6): 1554–1568.
42 Davidson F, Simmonds P, Ferguson JC *et al.* Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5′ non-coding region. *J Gen Virol* 1995; 76(5): 1197–1204.
43 Enomoto N, Sakuma I, Asahina Y *et al.* Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J Clin Invest* 1995; 96(1): 224–230.
44 Staden R, Beal KF, Bonfield JK. The Staden package, 1998. *Methods Mol Biol* 2000; 132: 115–130.
45 Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific

gap penalties and weight matrix choice. *Nucl Acids Res* 1994; 22: 4673–4680.
46 Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; 4: 406–425.
47 Posada D, Crandall KA. Selecting the best-fit model of nucleotide substitution. *Syst Biol* 2001; 50(4): 580–601.
48 Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Contr* 1974; 19: 716–723.
49 Swofford D. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Sunderland, MA: Sinauer Associates, 2002.
50 Rohlf FJ. Algorithm 76. Hierarchical clustering using the minimum spanning tree. *Computer J* 1973; 16: 93–95.
51 Schneider S, Roessli D, Excoffier L. Arlequin ver. 2000: A Software for Population Genetics Data Analysis. Switzerland: Genetics and Biometry Laboratory, University of Geneva, 2000.
52 Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 1992; 131: 479–491.
53 Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* 1986; 17: 57–86.
54 Posada D, Crandall KA. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). *Mol Biol Evol* 2001; 18(6): 897–906.
55 López-Labrador FX, Ampurdanés S, Giménez-Barcons M *et al.* Relationship of the genomic complexity of hepatitis C virus with liver disease severity and response to interferon in patients with chronic HCV genotype 1b infection. *Hepatology* 1999; 29(3): 897–903.
56 Sullivan DG, Wilson JJ, Carithers J, Perkins JD, Gretch DR. Multigene tracking of hepatitis C virus quasispecies after liver transplantation: correlation of genetic diversification in the envelope region with asymptomatic or mild disease patterns. *J Virol* 1998; 72(12): 10036–10043.
57 Duffy M, Salemi M, Sheehy N *et al.* Comparative rates of nucleotide sequence variation in the hypervariable region of E1/E2 and the NS5b region of hepatitis C virus in patients with a spectrum of liver disease resulting from a common source of infection. *Virology* 2002; 301: 354–361.
58 Berenguer M, Aguilera V, Prieto M *et al.* Delayed onset of severe hepatitis C-related liver damage following liver transplantation: a matter of concern? *Liver Transpl* 2003; 9(11): 1152–1158.
59 Chazouilleres O, Kim M, Combs C *et al.* Quantitation of hepatitis C virus RNA in liver transplant recipients. *Gastroenterology* 1994; 106(4): 994–999.