Compartmentalization of Hepatitis C Virus Variants in Patients With Hepatocellular Carcinoma

Paula S. Pérez,[1] Federico A. Di Lello,[1,2] Eduardo G. Mullen,[3] Omar A. Galdame,[3] Beatriz I. Livellara,[3] Adrián C. Gadano,[3] Rodolfo H. Campos,[1,2] and Diego M. Flichman[1,2]*

[1]*Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina*
[3]*Hospital Italiano de Buenos Aires, Buenos Aires, Argentina*

Chronic Hepatitis C Virus (HCV) infection is a major risk for hepatocellular carcinoma (HCC) development. HCV Core protein has been associated with the modulation of potentially oncogenic cellular processes and E2 protein has been useful in evolutive studies to analyze the diversity of HCV. Thus, the aim of this study was to evaluate HCV compartmentalization in tumoral, non-tumoral liver tissue and serum and to identify viral mutations potentially involved in carcinogenesis. Samples were obtained from four patients with HCC who underwent liver transplantation. Core and E2 were amplified, cloned and sequenced. Phylogenies and BaTS Test were performed to analyze viral compartmentalization and a signature sequence analysis was conducted by VESPA. The likelihood and Bayesian phylogenies showed a wide degree of compartmentalization in the different patients, ranging from total clustering to a more scattered pattern with small groups. Nevertheless, the association test showed compartmentalization for the three compartments and both viral regions tested in all the patients. Signature amino acid pattern supported the compartmentalization in three of the cases for E2 protein and in two of them for Core. Changes observed in Core included polymorphism R70Q/H previously associated with HCC. In conclusion, evidence of HCV compartmentalization in the liver of HCC patients was provided and further biological characterization of these variants may contribute to the understanding of carcinogenesis mediated by HCV infection.

Key words: hepatitis C virus; hepatocellular carcinoma; compartmentalization; liver transplant

INTRODUCTION

Chronic infection with hepatitis C virus (HCV) is considered a major risk for hepatocellular carcinoma (HCC) development [1]. According to recent estimates, over 185 million people are infected with HCV worldwide [2] and the rate of HCC progression among this population ranges from 1% to 3% over 30 years [3]. HCC is a leading cause of cancer-related death and its incidence is reported to be increasing in the last years [4].

Viral-induced HCC pathogenesis is related to the inflammation and regeneration environment that results from chronic liver damage. In addition, different HCV proteins are associated with the modulation of cellular processes with consequent potential oncogenic transformation [5]. In particular, Core protein has been reported to interact with several cellular proteins [6] and to modulate the control of cell proliferation and apoptosis [7–10].

Additionally, HCV is a highly variable virus which exists in the infected individual as a population of closely related variants, denominated quasispecies. The study of viral populations by analyzing the E2 region specially has contributed greatly to the understanding of viral evolution [11]. In this context, only selected HCV variants from the plethora of sequences present in an infected patient would be able to associate with specific cells. This phenomenon is described as compartmentalization, and has been observed in several studies [12–14].

Association of HCV heterogeneity with pathogenesis of HCC is a controversial issue. Molecular surveys have associated the presence of HCV variants harboring specific amino acid polymorphisms with a higher risk of HCC development [15–17]. However, most of these studies were conducted on serum or plasma

Abbreviations: AI, association index; BaTS, Bayesian tip-association significance; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; MC, maximum monophyletic clade; MCCT, maximum clade credibility trees; ML, maximum likelihood; NT, non-tumoral liver tissue; PS, parsimony score; RSCU, relative synonymous codon usage; S, serum; S_N, normalized Shannon entropy; T, tumoral liver tissue; VESPA, viral epidemiology signature pattern analysis.

Grant sponsor: Consejo Nacional de Investigaciones Científicas y Técnicas; Grant number: PIP 112-201101-00215; Grant sponsor: Agencia Nacional de Promoción Científica y Tecnológica; Grant number: PICT 2011-0705; Grant sponsor: Fundación Florencio Fiorini; Grant number: 2011-2012

*Correspondence to: Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956, 4° piso, (1113), Ciudad de Buenos Aires, Argentina.

Received 23 September 2015; Revised 3 April 2016; Accepted 2 May 2016

DOI 10.1002/mc.22500

Published online in Wiley Online Library (wileyonlinelibrary.com).

samples. Although circulating virus is thought to represent liver population, some studies have shown evidences that plasma variants differed from those in the liver [18,19]. Moreover, a few studies have found different viral variants according to tumoral or non-tumoral tissue sub-localization within the liver of HCC patients [20–25]. However, these studies analyzed only one protein and serum was not generally evaluated. In the present work, both serum and liver tumoral and non-tumoral compartments were included. Additionally, the analysis was based on two different structural proteins: Core, associated with viral pathogenesis, and E2, which provides suitable genetic information for variability evaluation.

Thus, the aim of this study was to evaluate HCV compartmentalization in tumoral, non-tumoral liver tissue and serum from patients with HCC, using different phylogenetic approaches, and to identify viral mutations potentially involved in carcinogenesis.

MATERIALS AND METHODS

Patients and Samples

Four patients infected with HCV subtype 1b, with cirrhosis and histologically proven HCC were included in this study. None of them presented co-infection with hepatitis B virus or human immunodeficiency virus. All the patients underwent liver transplantation for advanced liver disease. Fresh tumoral (T) and non-tumoral (NT) tissue samples were collected from the explanted liver and then stored at −80°C until use. Serum (S) samples were also obtained from the patients, either previous or post transplantation. Informed consent was obtained from each patient before the study was initiated.

RNA Extraction and Amplification

RNA was extracted from 200 μL of serum using the commercial kit High Pure Viral RNA Kit (Roche, Germany) according to manufacturer's protocol. Trizol reagent (Life Technologies, Carlsbad, CA) was used for RNA extraction from 50–100 mg of frozen liver samples, according to manufacturer's instructions.

Reverse transcription reaction was performed on 4.5 μL RNA using reverse transcriptase M-MLV (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and using the antisense primer EA (5′ RAA RCA RTG CGT GGG GCA 3′). A genomic fragment comprising Core, E1 and partial E2 (1740 bp) was amplified by nested PCR using Platinum Taq polymerase (Invitrogen) and the following external primers ES (5′ CGA AAG GCC TTG TGG TAC TG 3′, sense), EA (5′ RAA RCA RTG CGT GGG GCA 3′, antisense), and internal primers IS (5′ TTG TGG TAC TGC CTG ATA GGG T 3′), IA (5′ AAC CCC GTG TAG TTC ATC CA 3′). The cycling protocol was: denaturing at 94°C 5 min; followed by 40 cycles in the first round and 30 cycles in the second one of 94°C 40 s, 48°C 1 min and 72°C 3 min, increasing extension time 2 min every 10 cycles; and a final extension at 72°C for 5 min. For two of the samples for which the previous PCR was negative (T sample from patient 2 and NT sample from patient 4), Core and E2 proteins were amplified separately (593 bp each) by nested PCR using the primers: ES, EAcore (5′ GAC CGG YAY CCA RCA CCG AGA 3′, antisense), IS and IAcore (5′ GGA AGA TAG ARA ARG AGC AAC C 3′, antisense) for Core; and ESe2 (5′ TGG GAT AGA TGA TGA ACT GGT 3′, sense), EA, ISe2 (5′ TCC ATG GTG GGG AAC TGG GC 3′, sense) and IA for E2. The cycling protocol was: denaturing at 94°C 5 min; followed by 40 cycles in the first round and 25 cycles in the second one of 94°C 30 s, 48°C 30 s and 72°C 1 min for Core 1st round or 45 s in all other cases; and a final extension at 72°C for 5 min.

Cloning and Sequencing

Amplified fragments were cloned using pGEM-T Easy kit (Promega, Madison, WI) according to the manufacturer's conditions, which provides a blue/white system for clone selection. Thirty to fifty white colonies were screened by colony PCR for each sample. PCR products of positive clones were sequenced in both senses by an ABI automatic sequencer.

Phylogenies and Compartmentalization Test

The nucleotide sequences were aligned using Muscle v3.8.31 program and alignments were arranged using BioEdit v7.2.0. The most appropriate substitution model for each dataset was selected using jModelTest v2.1.3, assessed by the Akaike Information Criterion. Phylogenetic trees were inferred using the maximum likelihood (ML) method (PhyML 3.0) and their reliability was assessed by 1000 replicates of bootstrap resampling. For monophyly evaluation, sequences with high similarity to the clones were retrieved from GenBank using the BLAST tool and the phylogenetic analysis was evaluated by 100 replicates of bootstrap resampling.

The association between phylogeny and compartmentalization was statistically assessed in the Bayesian framework implemented in BaTS software. The input trees were obtained using BEAST package v1.8.0. All Monte Carlo Markov Chains were run for 10^8 generations in order to achieve an effective sample size >200 in all the parameters. A relaxed uncorrelated lognormal molecular clock with a fixed rate of 1.0 s/s/y and the Bayesian Skyride model were selected for the analysis due to their flexibility to accommodate different population dynamics. An input of 9000 trees was used for BaTS test, and three states were defined according to the compartment from which the sequences were obtained (serum, tumoral tissue, or non-tumoral tissue) in order to calculate three compartmentalization indexes: Parsimony Score (PS), Association Index (AI) and Maximum Monophyletic Clade (MC). The expected values for the indexes under the null hypothesis were

estimated by 1000 randomized sets. Maximum clade credibility trees (MCCT) were also obtained from the Bayesian runs previously described.

Diversity and Complexity of Quasispecies

The evolutionary nucleotide mean genetic distances within each group were estimated with MEGA6 using maximum Composite Likelihood method. Standard errors were obtained by bootstrap procedure (100 replicates). Genetic complexity was calculated with amino acid sequences and expressed as normalized Shannon entropy (S_N), being $S_N = \frac{S}{\ln N}$, where $S = -\Sigma_i(pi \times \ln pi)$ and N is the total number of clones in each sample. In this case, pi is the frequency of each clone in the quasispecies. S_N theoretically varies from 0 (only one variant detected) to 1 (maximum quasispecies complexity).

Signature Sequence Analysis

The VESPA software was used to search for signature patterns that distinguish tumoral liver tissue HCV variants from those in the corresponding non-tumoral tissue, both at the nucleotide and amino acid level. A 0.7 threshold was applied. The residues and nucleotides depicted by this analysis were verified by visual inspection of the alignments.

Epitope Analyses

Continuous antibody epitopes from Core and E2 consensus protein sequences from NT and T samples from each patient were predicted using on line IEDB Analysis Resource (http://tools.iedb.org/main/), by the BepiPred method [26]. This method uses a combination of a hidden Markov model and a propensity scale, based on parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity, and antigenic propensity of polypeptides chains, to predict the location of linear B-cell epitopes.

Effect of Nucleotide Substitutions on Codon Usage

Codon usage was evaluated for Core and E2 proteins from T and NT variants by calculating the Relative Synonymous Codon Usage (RSCU) index for each synonymous codon using DAMBE software v5.0.7. The RSCU value of a codon is its observed frequency divided by its expected frequency in the absence of usage bias. The correlation among RSCU values from T variants and NT variants was analyzed using the Spearman correlation test with the GraphPad software. In addition, to comparatively evaluate HCV and human codon usages, the correlation among RSCU values was analyzed using the most expressed liver genes [27].

Effect of Core Nucleotide Substitutions on RNA Structure

RNA secondary structural models were generated using on line software RNA Fold [28] by minimum free energy prediction. The inputs were the consensus Core sequences for each patient and sample (70% majority criterion), including a small fragment from the 5′UTR (19 nt). Already known RNA structures in Core [29] were set as constraints for the folding. Structures obtained for T and NT samples for each patient were compared by visual inspection, determining the effects of nucleotide changes on the structure.

Nucleotide Sequences Accession Numbers

Nucleotide sequences for Core and E2 clones have been deposited in GenBank under accession numbers KT709951–KT710160 and KT709747–KT709950, respectively.

RESULTS

The characteristics of the four patients analyzed are expressed in Table 1. A total of 414 cloned sequences were obtained and an average of 17 (11–24) sequences

Table 1. Characteristics of Patients and Number of Clones Analyzed

Patient	Age	Gender	METAVIR score	HCC Edmondson–Steiner grade	Viremia ($\log_{10}$ copies/ml)	Tissue	Number of clones analyzed
1	59	F	A2F4	III	5.71	NT	19
						T	12
						S[#]	18
2	67	F	A2F4	II	6.51	NT	18
						T	24, 20[a]
						S[#]	11
3	65	F	A2F4	II	6.67	NT	16
						T	17
						S[§]	13
4	55	M	A2F4	II	6.51	NT	21, 19[a]
						T	17
						S[§]	24

T, tumoral; NT, non-tumoral; S[#], serum collected pre-transplant; S[§], Serum collected post-transplant.
[a]For Core and E2 regions, respectively.

from each sample for both of the genomic regions analyzed (Table 1).

Phylogenetic Analyses

The phylogenetic trees obtained for each of the patients analyzed showed a distinct topology pattern regarding clustering of variants from different localizations (Figures 1 and 2). For the samples for which the 1740 bp fragment was cloned and sequenced a conjunct interpretation of Core and E2 phylogenies was done.

Firstly, for patient 1 (Figure 1a), clones from each sample: T, NT, and S, were grouped in highly supported clades (>70% bootstrap value) for E2 and Core proteins. Trees for both regions showed the same topology. NT sequences were represented by two lineages, one of them poorly sampled and distant from the main one. For patient 2 (Figure 1b), sequences from T tissue showed several clusters, while NT and S clones were represented by one lineage. This result was observed in both regions analyzed. For patient 3 (Figure 2a), the E2 tree revealed the existence of two lineages in the liver. Interestingly, each of them was formed by two monophyletic sub-lineages corresponding to T or NT clones exclusively; however, this clustering was not observed in the Core region. All serum sequences grouped separately from liver samples in both analyses. Finally, for patient 4, small clusters of clones from each sample were observed, although the topology support was limited. In addition, Bayesian phylogenies were performed for the four cases under study, and the MCCTs were obtained (Supplementary Figure S1). In all the cases, tree topologies were consistent with those obtained by ML method.

Common origin of clones was evaluated by E2 region, since it gives appropriate phylogenetic signal. The likelihood tree obtained showed that clones from each patient formed monophyletic clusters with statistic support (data not shown).

Compartmentalization Test

In all the cases analyzed, the phylogenies obtained for Core and E2 showed global evidence of compartmentalization assessed by the AI and the PS ($P < 0.001$) (Supplementary Table S1). In addition, MC index analysis, which is a statistic for each particular trait, showed that the association was statistically supported ($P < 0.05$) for the three compartments and both regions tested in all the patients.

Diversity and Complexity of Quasispecies

Genetic diversity and complexity results for each patient are depicted in Table 2. No significant differences were observed between T and NT compartments. However, entropy values for Core protein T samples ranged 0.424–0.814 (Mean = 0.611), whereas the entropy for NT samples ranged 0.356–0.523 (Mean = 0.427). Thus, quasispecies complexity showed a clear tendency to be higher within the T samples than for NT samples ($P = 0.055$) for Core but not for E2 protein.

Amino Acid and Nucleotide Patterns

The individual analysis on Core sequences showed amino acids exclusive from one compartment for patients 1, 2, and 3 (Figure 3). Moreover, differences between T and NT variants were seen in cases 1 and 2. In the E2 protein analysis, signature amino acids for each compartment were also identified for cases 1, 2, and 3 (Supplementary Figure S2), most of them located within the hypervariable regions in E2. For patient 3, different amino acid residues between T and NT tissue variants were identified in both of the viral lineages observed in the phylogeny (L1 and L2). In addition, nucleotide analysis on both regions revealed synonymous differences between T and NT variants in all the patients analyzed. On the other hand, no differential signature pattern between T and NT variants was observed when sequences were analyzed altogether (data not shown).

Effect of Amino Acid Substitutions on Predicted Epitopes

Differential positions observed in Core and E2 proteins between NT and T were mainly located within a predicted epitope (4/5 in Core, and 8/14 in E2). Locations of the predicted epitopes were the same for NT and T sequences, although lengths varied slightly for E2 protein in patient 1. In general, the change in the amino acid residue did not lead to disruption or creation of an epitope, with the exceptions of substitutions from T samples: 190F in Core from patient 1 and 405Q in E2 from patient 2, which disrupted the epitope predicted for the NT sample at that position.

Effect of Nucleotide Substitutions on Codon Usage

In order to evaluate the existence of codon usage bias by the viral populations in each compartment, codon frequencies from T and NT sequences were compared. Codon usage of T variants did not result significantly different from that of NT variants for Core and E2 proteins in all the four cases. Additionally, codon usage of these viral populations showed correlation with human codon usage of liver-specific highly expressed genes in all the cases, except for the E2 region from patient 2. On the other hand, when codons were analyzed individually, substitutions observed in T tissue tended to favor the usage of more prevalent codons than the ones observed in NT (data not shown).

Effect of Core Nucleotide Substitutions on RNA Structure

RNA structures obtained for sequences derived from T samples were not different from the corresponding NT samples in patients 2, 3, and 4. For patient 1, modeled RNA structures for T and NT sequences differed at the 3′ end, although this alteration was

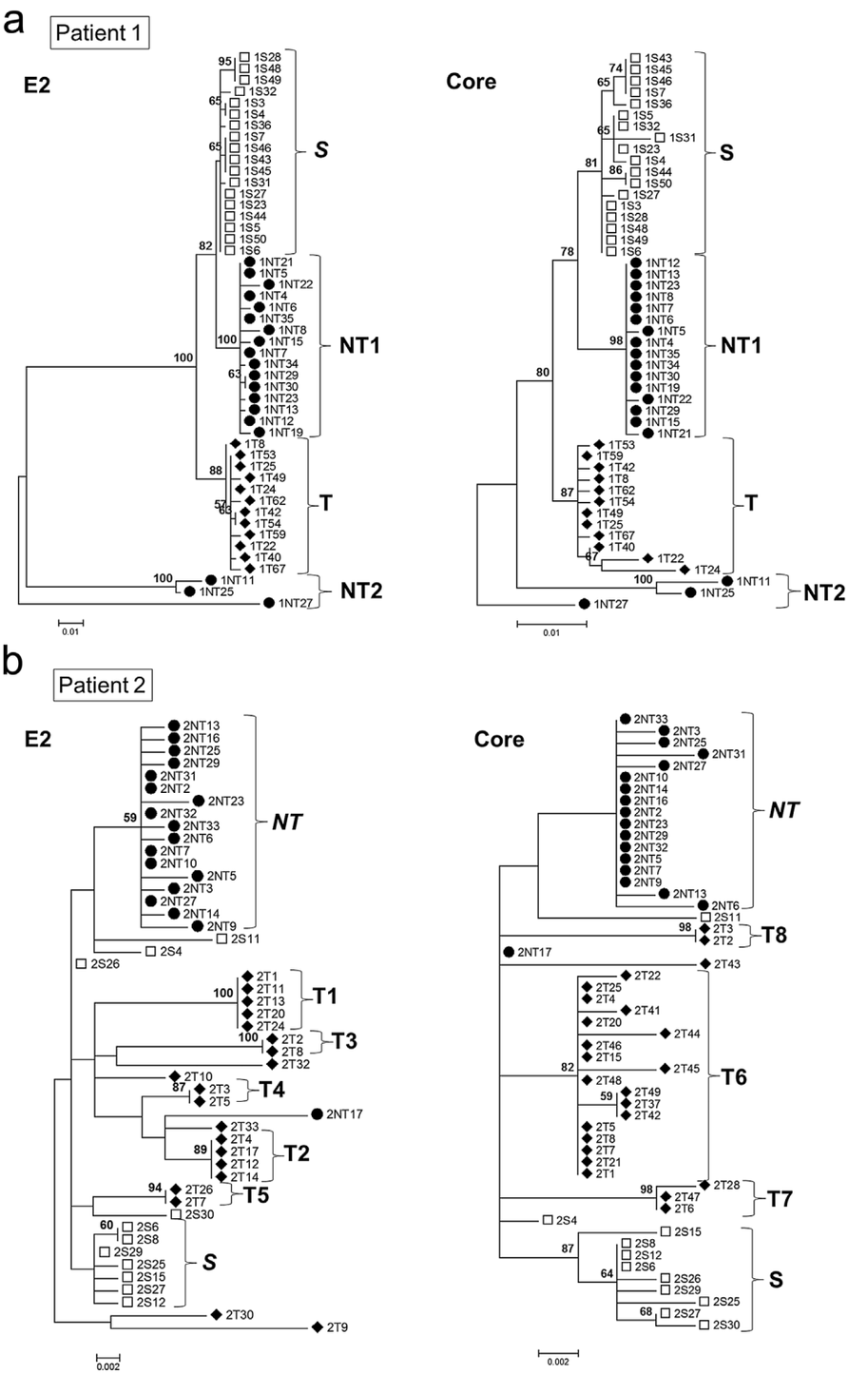

Figure 1. Likelihood phylogenetic trees for E2 and Core corresponding to: a) Patient 1 and b) Patient 2. Taxa are named according to patient, compartment (T, tumoral; NT, non-tumoral; S, serum) and clone number. Shapes of the symbols represent compartment (T, shaded diamond; NT, shaded circle, S, empty square). The numbers above the branches represent the bootstrap proportion (over 1000 pseudoreplica) over that 50%. Highlighted groups indicate the lineages observed. Scale bar represents substitutions per site.

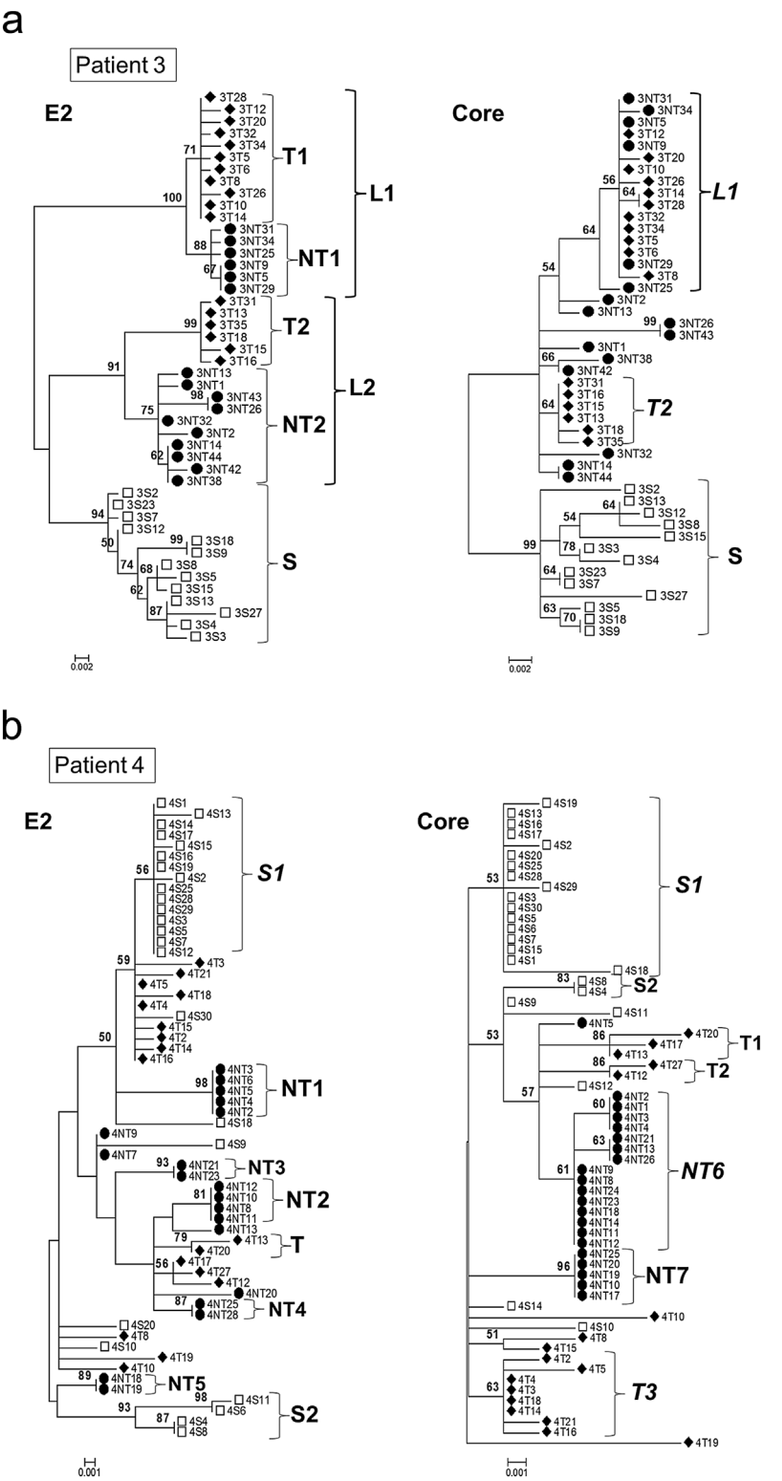

Figure 2. Likelihood phylogenetic trees for E2 and Core corresponding to: a) Patient 3 and b) Patient 4. Taxa are named according to patient, compartment (T, tumoral; NT, non-tumoral; S, serum) and clone number. Shapes of the symbols represent compartment (T, shaded diamond; NT, shaded circle; S, empty square). The numbers above the branches represent the bootstrap proportion (over 1000 pseudoreplica) over that 50%. Highlighted groups indicate the lineages observed. Scale bar represents substitutions per site.

Table 2. Diversity and Complexity of Quasispecies Within Compartments

	Core						E2					
	Nucleotide distance Mean (±SE)			Shannon's entropy (S_N)			Nucleotide distance Mean (±SE)			Shannon's entropy (S_N)		
Patient	NT	T	S	NT	T	S	NT	T	S	NT	T	S
1	0.012 (±0.004)	0.006 (±0.002)	0.004 (±0.001)	0.404	0.814	0.606	0.069 (±0.040)	0.004 (±0.001	0.004 (±0.001)	0.757	0.639	0.484
2	0.002 (±0.001)	0.007 (±0.002)	0.008 (±0.002)	0.427	0.550	0.790	0.005 (±0.001)	0.020 (±0.004)	0.008 (±0.002)	0.493	0.640	0.695
3	0.008 (±0.002)	0.006 (±0.002)	0.010 (±0.002)	0.523	0.424	0.636	0.026 (±0.007)	0.020 (±0.005)	0.009 (±0.002)	0.789	0.816	0.875
4	0.006 (±0.002)	0.011 (±0.002)	0.004 (±0.001)	0.356	0.655	0.215	0.014 (±0.003)	0.015 (±0.003)	0.013 (±0.003)	0.564	0.816	0.357

T, tumoral; NT, non-tumoral; S, serum.

minor and located outside any described elements (data not shown).

DISCUSSION

In this study, we characterized viral populations from liver tumoral and non-tumoral tissue, and serum samples from patients with HCC. Our results showed compartmentalization between liver and serum, and also between tumoral and non-tumoral hepatocytes. Moreover, we identified amino acid substitutions in the tumoral sequences exclusive from this compartment.

The likelihood and Bayesian phylogenies inferred for the four cases analyzed showed a wide degree of compartmentalization, ranging from total clustering observed in patient 1 to a more scattered pattern with small groups in the other patients. There is a paucity of information regarding HCV populations in tumoral and non-tumoral tissue in HCC patients. Our results are in line with the study performed by Sobesky et al. who analyzed viral population of Core region in 7 patients and determined compartmentalization of T and NT variants by Mantel's Test [20]. In contrast, Rüster et al. carried out a study in 8 patients and, although the majority of sequences were found exclusively either in one or other tissue, they did not find a strict phylogenetic separation of T and NT Core variants by neighbor-joining method [22]. It is worth noting that these studies were performed on conserved proteins and only by distance matrix based methods.

In the present work, phylogenetic clustering of variants from each compartment was present in more or less extent in all cases and the Bayesian test results supported the association between compartment and phylogeny. Additionally, increased quasispecies complexity was observed in T tissue in three out of four patients for both regions. Overall analysis comparing T versus NT entropies revealed a tendency of higher genetic complexity within the Core region of T tissue isolates, but not in E2. Although significant differences were not reached, probably due to the low number of patients, differences depicted cannot be ignored. Moreover, this finding is in agreement with what was previously stated by Rüster et al. [22]. These results suggest that compartmentalization of HCV variants would exist within the liver of HCC patients between T and NT tissue. In particular, compartmentalization was established in this study for two of the structural proteins and on the basis of more robust phylogenetic and association methods than neighbor joining approaches.

Compartmentalization observed between T and NT HCV populations may be explained at least by two non-exclusive hypotheses. On the one hand, differences observed between variants may correspond to molecular adaptation of the virus either to T or NT cells, and selective pressure imposed by each micro-

	7	12	16	37	49	70	75	91	101	110	147	149	150	157	161	187	190
C1b	P	K	N	L	T	**q**	**a**	m	R	**T**	A	R	A	V	G	I	S
1NT	.	.	.	**M**	P	R	T	.	.	S	.	.	.	.	.	V	.
1T	**L**	.	.	.	P	R	T	.	.	S	.	.	.	.	.	V	**F**
1S	.	.	**S**	.	P	R	T	.	.	S	.	.	.	.	.	V	.
2NT	.	.	.	.	.	.	.	.	**H**	S	V	K	V	A	.		
2T	.	.	.	.	.	**R**	.	.	.	S	V	K	V	A	.		
2S	.	.	.	.	.	.	T	.	.	S	V	K	V	A	.		
3NT	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
3T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
3S	.	**N**	.	.	.	.	T	.	.	.	.	.	**V**	.	.	.	.
4NT	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.		
4T	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.		
4S	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.		

Figure 3. Core amino acid signatures for serum, tumoral and non-tumoral tissues. The alignment shows amino acids for each sample found to be different from a consensus HCVA1b Core (C1b) constructed from 100 sequences retrieved from GenBank with a 70% majority criteria (lower case indicates <70% majority). Changes in positions 12, 70, 91, and 161 were previously associated with HCC. Polymorphic sites (Hx > 0.5) were determined by entropy calculation and are indicated in bold on C1b. Consensus sequences of samples are named after number of patient and compartment (T, tumoral; NT, non-tumoral; S, serum). Differences between T and NT are shaded. Amino acids exclusive from one compartment are bold-underlined. Amino acids are indicated by singleAletter code and those identical to C1b are represented with dots. Positions are related to H77 polyprotein start (Acc. Number AF009606).

environment may lead to fixation of specific mutations in each compartment. On the other hand, HCV behavior as quasispecies implies that many variants coexist within an infected individual and are transmitted as a complex population. Some of these variants may carry biological characteristics which enhance carcinogenesis. Thus, development of carcinoma nodules could possibly occur more likely in the cells infected with these variants.

Intriguingly, viral populations from serum were not represented in T or NT tissues. In particular, a clear serum compartmentalization was observed for the patients having pre-transplant samples. This result is in accordance with other studies which showed different viral populations between serum and liver [14,18,19,25]. This finding could be attributed, among other causes, to: (i) the existence of slow replication or defective variants in the liver that would be less represented among circulating variants, (ii) a difference in HCV variant clearance rates, and/or (iii) the contribution of the extrahepatic viral replication. In this context, even when serum samples are the best choice for prediction, screening and diagnosis, analysis on liver viral variants could add a piece of knowledge on the investigation of molecular mechanisms involved in HCC development by HCV infection.

The finding of viral compartmentalization in the tissues analyzed suggests that signature molecular differences may exist between T and NT compartments. At the amino acid level, a signature pattern was found in three of the cases for E2 protein and in two of them for Core. Substitutions were mainly located within predicted antibody epitopes, and the change in the amino acid residue did not disrupt the epitope in most cases. In addition, changes observed for E2 protein were found mostly in the hypervariable regions. Therefore, it is possible that E2 substitutions correspond to differential selection pressures exerted by the immune system in each compartment. In the Core protein, most of the signature amino acids were located in domain 1, while one residue was part of the signal peptide. Among the observed substitutions in Core, polymorphism R70Q/H has been previously associated with HCC [15,30], and it has been shown that the ratio of mutant residues 70Q/H increased as liver disease advanced to HCC [17]. In our study, 70Q variant was found both in serum and liver tissues for three patients, one of whom presented a mixed population in T tissue where 70R and 70Q variants accounted 75% and 25% of the population respectively. Interestingly, when compared to 100 HCV-1b Core sequences randomly selected from GenBank, tumoral substitutions observed in patient 1 (P7L, S190F) were not found. Besides, substitution P7L has been previously reported in Core variants isolated from tumoral tissue which inhibited TGF-β pathway [24], thus correlating HCV genomic variability with a gain in a biological function associated with carcinogenesis. While in contrast, the remaining changes observed had not been previously described. Since Core is known to modulate cellular processes, substitutions in this protein could play a differential role in this regulation.

In addition to amino acidic changes, compartmentalization was also represented at the nucleotide level by synonymous substitutions, which accounted for the majority of substitutions. It is known that codon usage has been described as a viral mechanism to regulate its expression [31]. Although nucleotide sequences from T and NT showed a correlation in codon frequencies, both among themselves and with

genes from human liver, it is worth noting that, when codons were analyzed individually, substitutions observed in T tissue tended to favor the usage of more frequent codons than the ones observed in NT. On the other hand, even though nucleotide substitutions in the Core region could alter predicted secondary RNA structures which are believed to be related to alternative translation and viral assembly [29,32,33], this effect was not observed in our samples.

In spite of the limitation given by the number of patients included in this study, it is usual to carry out studies with reduced cohorts due to the inherent difficulty to obtain liver samples; therefore, results obtained even from a small number of patients are valuable. In order to increase the sample size, signature pattern analysis was performed on Sobesky et al. dataset [20]. Although substitutions characteristic to tumoral tissue where found in three out of seven patients, they all differed from the ones observed in this study. This observation highlights the fact that even when more patients are added, inter-individual variations are still observed. The sequencing of a limited number of variants may also be a limitation. However, they were sufficient for determining compartmentalization and the main molecular differences between each compartment. Moreover, a previous study has shown that sequencing 10 viral variants is sufficient for quasispecies evaluation [34]. Although the use of deep sequencing techniques would produce an increase in the observed diversity [35], it would not affect the qualitative results of this work.

In conclusion, the present study provided evidence of HCV compartmentalization within the liver of patients with HCC, showing that HCV variants differed between tumoral and non-tumoral tissue. Further characterization of these variants by means of biological experiments may contribute to the understanding of carcinogenesis mediated by HCV infection.

ACKNOWLEDGMENTS

This work was supported by Grants from the Consejo Nacional de Investigaciones Científicas y Técnicas, PIP 112-201101-00215, Agencia Nacional de Promoción Científica y Tecnológica, PICT 2011–0705 and Fundación Florencio Fiorini 2011–2012.

REFERENCES

1. de Martel C, Maucort-Boulch D, Plummer M, Franceschi S. Worldwide relative contribution of hepatitis B and C viruses in hepatocellular carcinoma. Hepatology 2015;62:1190–1200.
2. Mohd Hanafiah K, Groeger J, Flaxman AD, Wiersma ST. Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to HCV seroprevalence. Hepatology 2013;57:1333–1342.
3. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology 2012;142:1264–73.e1.
4. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015;65: 87–108.
5. Banerjee A, Ray RB, Ray R. Oncogenic potential of hepatitis C virus proteins. Viruses 2010;2108–2133.
6. de Chassey B, Navratil V, Tafforeau L, et al. Hepatitis C virus infection protein network. Mol Syst Biol 2008;4:230.
7. Higgs MR, Lerat H, Pawlotsky J-M. Hepatitis C virus-induced activation of b-catenin promotes c-Myc expression and a cascade of pro-carcinogenetic events. Oncogene 2013;32: 4683–4693.
8. Ngo HTT, Pham LV, Kim J-W, Lim Y-S, Hwang SB. Modulation of mitogen-activated protein kinase-activated protein kinase 3 by hepatitis C virus core protein. J Virol 2013;87: 5718–5731.
9. Jahan S, Ashfaq U a, Khaliq S, Samreen B, Afzal N. Dual behavior of HCV Core gene in regulation of apoptosis is important in progression of HCC. Infect Genet Evol 2012;12: 236–239.
10. Koike K. Hepatitis C virus contributes to hepatocarcinogenesis by modulating metabolic and intracellular signaling pathways. J Gastroenterol Hepatol (Australia) 2007;22 Suppl 1: S108–S111.
11. Culasso ACA, Baré P, Aloisi N, Monzani MC, Corti M, Campos RH. Intra-host evolution of multiple genotypes of hepatitis C virus in a chronically infected patient with HIV along a 13-year follow-up period. Virology 2014;449:317–327.
12. Gismondi MI, Díaz Carrasco JM, Valva P, et al. Dynamic changes in viral population structure and compartmentalization during chronic hepatitis C virus infection in children. Virology 2013;447:187–196.
13. Vera-Otarola J, Barría MI, León U, et al. Hepatitis C virus quasispecies in plasma and peripheral blood mononuclear cells of treatment naïve chronically infected patients. J Viral Hepat 2009;16:633–643.
14. Ramirez S, Perez-Del-Pulgar S, Carrion J a, et al. Hepatitis C virus compartmentalization and infection recurrence after liver transplantation. Am J Transplant 2009;9:1591–1601.
15. Fishman SL, Factor SH, Balestrieri C, et al. Mutations in the hepatitis C virus core gene are associated with advanced liver disease and hepatocellular carcinoma. Clin Cancer Res 2009;15:3205–3213.
16. El-Shamy A, Shindo M, Shoji I, Deng L, Okuno T, Hotta H. Polymorphisms of the core, NS3, and NS5A proteins of hepatitis C virus genotype 1b associate With development of hepatocellular carcinoma. Hepatology 2013;58:555–563.
17. Miura M, Maekawa S, Takano S, et al. Deep-sequencing analysis of the association between the quasispecies nature of the hepatitis C virus core region and disease progression. J Virol 2013;87:12541–12551.
18. Blackard JT, Ma G, Sengupta S, et al. Evidence of distinct populations of hepatitis C virus in the liver and plasma of patients co-infected with HIV and HCV. J Med Virol 2014;86: 1332–1341.
19. Jouvencel A-C, Neau D, Faure M, et al. Plasma and liver hepatitis C virus variability in patients coinfected with human immunodeficiency virus. J Clin Microbiol 2006;44:1877–1880.
20. Sobesky R, Feray C, Rimlinger F, et al. Distinct hepatitis C virus core and F protein quasispecies in tumoral and nontumoral hepatocytes isolated via microdissection. Hepatology 2007; 46:1704–1712.
21. Alam SS, Nakamura T, Naganuma A, et al. Hepatitis C virus quasispecies in cancerous and noncancerous hepatic lesions: The core protein-encoding region. Acta Med Okayama 1999; 53:141–147.
22. Rüster B, Zeuzem S, Krump-Konvalinkova V, et al. Comparative sequence analysis of the core- and NS5-region of hepatitis C virus from tumor and adjacent non-tumor tissue. J Med Virol 2001;63:128–134.

23. Ogata S, Nagano-Fujii M, Ku Y, Yoon S, Hotta H. Comparative sequence analysis of the core protein and its frameshift product, the F protein, of hepatitis C virus subtype 1b strains obtained from patients with and without hepatocellular carcinoma. J Clin Microbiol 2002;40:3625–3630.
24. Pavio N, Battaglia S, Boucreux D, et al. Hepatitis C virus core variants isolated from liver tumor but not from adjacent non-tumor tissue interact with Smad3 and inhibit the TGF-beta pathway. Oncogene 2005;24:6119–6132.
25. De Mitri MS, Mele L, Chen CH, et al. Comparison of serum and liver hepatitis C virus quasispecies in HCV-related hepatocellular carcinoma. J Hepatol 1998;29:887–892.
26. Larsen JEP, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome Res 2006;2:2.
27. Plotkin JB, Robins H, Levine AJ. Tissue-specific codon usage and the expression of human genes. Proc Natl Acad Sci USA 2004;101:12588–12591.
28. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res 2008;36:70–74.
29. Tuplin A, Evans DJ, Simmonds P. Detailed mapping of RNA secondary structures in core and NS5B-encoding region sequences of hepatitis C virus by RNase cleavage and novel bioinformatic prediction methods. J Gen Virol 2004;85: 3037–3047.
30. Akuta N, Suzuki F, Kawamura Y, et al. Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. Hepatology 2007;46:1357–1364.
31. Pintó RM, Aragonès L, Costafreda MI, Ribes E, Bosch A. Codon usage and replicative strategies of hepatitis A virus. Virus Res 2007;127:158–163.
32. Mauger DM, Golden M, Yamane D, et al. Functionally conserved architecture of hepatitis C virus RNA genomes. Proc Natl Acad Sci 2015;112:3692–3697.
33. Pang PS, Elazar M, Pham E a, Glenn JS. Simplified RNA secondary structure mapping by automation of SHAPE data analysis. Nucleic Acids Res 2011;39:1–11.
34. Marukian S, Jones CT, Andrus L, et al. Cell culture-produced hepatitis C virus does not infect peripheral blood mononuclear cells. Hepatology 2008;48:1843–1850.
35. Gong L, Han Y, Chen L, et al. Comparison of next-generation sequencing and clone-based sequencing in analysis of hepatitis b virus reverse transcriptase Quasispecies heterogeneity. J Clin Microbiol 2013;51:4087–4094.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.