This is a postprint of an article published in
Gismondi, M.I., Becker, P.D., Díaz Carrasco, J.M., Guzman, C.A., Campos, R.H., Preciado, M.V.
Evolution of hepatitis C virus hypervariable region 1 in immunocompetent children born to HCV-infected mothers
(2009) Journal of Viral Hepatitis, 16 (5), pp. 332-339.

Evolution of hepatitis C virus (HCV) hypervariable region 1 in immunocompetent children born to HCV-infected mothers

María I. Gismondi[1], Pablo D. Becker[3], Juan M. Díaz Carrasco[1], Carlos A. Guzmán[3], Rodolfo H. Campos[2], María V. Preciado[1]

[1]Laboratorio de Biología Molecular, División Patología, Hospital de Niños Ricardo Gutiérrez; [2],Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Argentina; [3]Abteilung Vakzinologie, Helmholtz Zentrum für Infektionsforschung, Braunschweig, Deutschland.

Running head: HVR1 conservation in HCV-infected children

Correspondence to:

Dr. María Victoria Preciado, PhD; Laboratorio de Biología Molecular, División Patología, Hospital de Niños Ricardo Gutiérrez; Gallo 1330 (1425) Ciudad de Buenos Aires, Argentina; Pho: +54 11 4962 9138; Fax: +54 11 4963 4122; e-mail: preciado@conicet.gov.ar

Summary

Hepatitis C virus (HCV) hypervariable region 1 (HVR1) is the most variable region of the viral genome and its heterogeneity reflects the virus-host interplay during chronicity. Paediatric HCV-infected patients develop a liver disease with typical clinical features. Here, the evolution of HVR1 and its adjacent regions was ascertained in plasma samples of two HCV-positive children during a 5-year follow-up. We report an almost complete conservation of HVR1 sequence over time, with underlying nucleotide variability both within and outside HVR1, suggesting some kind of constraint on virus evolution, particularly within HVR1. Although overall d_N/d_S rates were <1 in both patients, a high resolution analysis of selection pressures exerted at the codon level revealed few sites subject to selection and an absolute predominance of invariable positions within HVR1. The HVR1 amino acid sequences showed the antigenic properties expected for this region. Taken together, these data suggest a peculiar evolutionary dynamics in our patients, which could be attributed to a mechanism of nucleotide invariability along with purifying selection operating on the HVR1. The lack of HVR1 variability may reflect the adaptation of the virus to a particular environment within each patient or a phenomenon of immune tolerance generated in our immunocompetent patients early in life.

Key words: Hepatitis C virus, HVR1, intra-host evolution, perinatal infection, conservation.

Introduction

Hepatitis C virus (HCV), an enveloped RNA virus belonging to the genus Hepacivirus in the *Flaviviridae* family, is the major cause of non-A non-B hepatitis worldwide (1). Chronic HCV infection may lead to cirrhosis and/or hepatocellular carcinoma; in fact, it represents the main indication for liver transplantation among adults (1). However, disease progression seems to be different in children, since they rarely develop severe hepatitis or cirrhosis. Children with chronic HCV infection show milder liver disease than adults, albeit displaying high HCV viral load (2, 3).

HCV displays a high degree of diversity, both at the nucleotide and at the amino acid level (4). In an infected individual, the virus exists as a mixture of distinct but closely related variants, which together are called quasispecies (5). The evolution of HCV quasispecies is modulated by viral and host factors. The former include the error-prone HCV RNA-dependent RNA polymerase, which generates an estimated mutation rate of 10^{-5} mutations/nucleotide/replication (6) and a replication rate of up to 10^{12} virions/day (7). Among the latter, the immune response against HCV constitutes a selective pressure for viral evolution. In particular, the hypervariable region 1 (HVR1), located at the N-terminus of HCV glycoprotein E2, is appropriate for studying HCV quasispecies dynamics, since its putative antigenic role makes it the most variable region of the viral genome. Despite this high variability, its conformation and physicochemical properties are conserved (8, 9).

The study of virus evolution constitutes a key matter for the unraveling of virus-host interactions. It becomes particularly important in chronic infections like HCV-related hepatitis, since the success of the infection depends on a subtle balance between virus replication and its control by the host's environment. The evolution of HCV quasispecies in a single patient has been extensively studied, both in treated and non-

treated adults (10-14). However, little is known about HCV evolution in chronically infected children, a clinical situation in which there is a considerable paucity of knowledge and even contrasting results. In the setting of perinatal HCV infection, HVR1 stability during the first weeks or months of life has been reported (15, 16), with the host's selective pressure being a major determinant in driving HCV evolution (15). Conversely, other authors showed increasing HVR1 diversification over time in chronically HCV infected children (17), and a high level of HVR1 diversification during the first year of life in perinatally HCV-HIV co-infected children (18).

Given that chronic HCV infection seems to have a different course in children than in adults, the study of virus evolution in the former is a matter of great interest. It seems plausible that, in the setting of paediatric infection, the virus may establish and maintain chronicity in a different way than the one described for chronic HCV infection acquired in adulthood. Considering that the host's environment modulates viral evolution through the many selective forces that operate in it, HVR1 diversification and evolution may display typical features in children. It is still under study whether these evolutionary events are related to disease progression and fibrosis score. Our aim was to perform a retrospective analysis of HVR1 evolution in two paediatric patients with chronic HCV infection during a prolonged follow-up.

Methods

Patients and samples

Two patients (ages 3 and 2 years at the beginning of this study) born to HCV-infected mothers (one HIV-HCV co-infected mother) were included. Both patients were infected by HCV genotype 1a/c, and they were negative for hepatitis B surface antigen and anti-human immunodeficiency virus. Blood samples of both patients were taken at

different time points during a follow-up of 60 and 63 months for patients 1 and 2, respectively. The biochemical and virological parameters examined during follow-up are summarized in Table 1. Institutional guidelines regarding human experimentation were followed.

Reverse transcription and PCR amplification

RNA was extracted from 200 µl plasma with Trizol reagent (Invitrogen) according to manufacturer's instructions. RNA was denaturated for 5 min at 70º C, and the reverse transcription reaction was performed for 60 min at 42º C using M-MLV reverse transcriptase (Promega) and 4 pmol antisense primer E2OA (5' TCTCAGGACAGCCTGAAGKGTTGAA 3') in a final volume of 10 µL. After heat inactivation at 95º C for 5 min and chilling on ice, the cDNA was amplified. The amplification was carried out using primers E2OS (5' GCCATATAACGGGTCACCGCATGGC 3', sense) and E2OA (antisense) in the first round, and E2NS (5' GGATATGATGATGAACTGGTC 3', sense) and E2NA (5' GGTGTTGAGGCTATCATTGCARTT 3', antisense), in the nested reaction. The PCR reactions consisted of 35 cycles of denaturation at 94º C for 15 s, annealing at 53º C for 30 s and polymerization at 68º C for 40 s. Reactions were carried out with *Pfx* DNA polymerase (Invitrogen). The 351 bp-amplicons (nucleotides 1296 to 1646 of HCV-genotype 1 prototypic secuence, GenBank accession number AF09069) were purified from agarose gel with QIAEXII gel extraction system (QIAGEN) according to manufacturer's instructions.

Cloning of PCR products

PCR products were subjected to the addition of 3'A-overhangs using *Taq* DNA polymerase (Invitrogen), cloned in a pCR2.1 vector using Topo TA cloning® Kit (Invitrogen), and transformed into *Escherichia coli* DH5α-competent cells. Plasmid

DNA from transformants producing white colonies was extracted from minipreps cultured overnight by alkaline lysis (19). Purified amplicons and purified recombinant plasmid DNA were sequenced bidirectionally in an Applied Biosystems Division Automated 3700 DNA Analyzer, using the ABI PRISM® BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer Applied Biosystems).

Sequence analysis

Nucleotide sequences corresponding to each patient were aligned using CLUSTALX program (20). Given the high variability expected within HVR1, the most appropriate model of evolution was inferred separately for this segment (81 bp) and its adjacent regions (AR, 220 bp) using Modeltest v.3.7 (21). Median genetic distances between all possible pairs of sequences both within and between samples of the same individual were calculated with the previously defined evolutionary model using PAUP* v.4.0.b10 (22).

The rates of non-synonymous nucleotide substitutions per non-synonymous site (d_N) and synonymous nucleotide substitutions per synonymous site (d_S) were calculated by the method of Nei and Gojobori (23) with Jukes Cantor correction, as implemented in the MEGA3 package (24).

Selection often occurs on a few amino acid positions, and so the mean d_N/d_S ratio over the entire sequence may lead to erroneous conclusions about the mechanisms underlying viral evolution. For this reason, positively and negatively selected sites were identified by a maximum likelihood method as implemented in HyPhy software (25). Briefly, an MGx94xHKY85 codon frequency model with discretized synonymous (α) and non-synonymous (β) per site substitution rates was assumed, and an empirical Bayes approach was employed to detect codons subject to positive or negative selection.

Amino acid sequences were deduced from the corresponding nucleotide sequences using BioEdit Software version 7.0.4.1 (26). Antigenicity was predicted from the amino acid sequences using Parker's method as implemented in the ANTHEPROT software v.6.0 (27).

Statistical analysis

Comparison of median intra- and inter-sample genetic distances was performed using Mann-Whitney or Kruskal-Wallis Tests (with Dunn's Multiple Comparisons Post-Test when appropriate) for non-parametric variables using GraphPad Instat Software, version 3.05. A p-value <0.05 was considered statistically significant.

Results

In a first attempt to study HCV evolution in perinatally infected children, amplification fragments encompassing the most variable region (HVR1) of the viral genome from consecutive plasma samples obtained during a prolonged follow-up were sequenced. Notably, the aligned HVR1 sequences exceptionally contained ambiguous positions (data not shown) in different samples from both patients, suggesting a strong predominance of a major HCV variant detected over other minor viral variants which may potentially be present in patient's samples. Moreover, the HVR1 nucleotide sequences did not change overall in consecutive plasma samples of both patients (data not shown).

To better evaluate HVR1 evolution, analysis of quasispecies dynamics becomes mandatory. The sequencing of molecular? clones of the amplified region is a valuable approach for the description of the complexity and diversity of the quasispecies during the course of chronic HCV infection. Thus, in order to characterize the underlying quasispecies behavior of the apparently stable HVR1, molecular? clones of the

amplified region of some plasma samples were obtained and sequenced (Table 1). The distributions of genetic distances both within and between samples (taken as a measure of quasispecies diversity) are summarized in Table 2. The intra-sample analysis displayed low degree of diversity revealing the existence of more than one HVR1 variant in each sample of both patients. In patient 1, intra-sample HVR1 diversity decreased continuously and significantly between samples taken at 0 and 47 months ($p<0.001$), with a subsequent increase around 61 months of follow-up ($p<0.001$). Meanwhile, HVR1 intra-sample diversity showed no significant difference between samples of patient 2 ($p=0.5370$). In the inter-sample analysis, patient 1 showed increasing diversification of HVR1 over time ($p<0.05$), whereas patient 2 displayed a significant reduction in the median genetic distance between the sample taken at the beginning of the study and subsequent samples ($p=0.012$). These results indicate that the apparently homogeneous HVR1 sequences detected initially showed in fact a quasispecies distribution, which in turn evolved with typical features in each patient.

Upon translation of nucleotide sequences, both patients displayed an almost complete conservation of the HVR1 amino acid sequence throughout follow-up, as shown by a high number of identical clones (data not shown). The most frequent variant identified at the beginning of the study (time point month 0) always predominated during follow-up (Table 3). Noteworthy, the minor HVR1 variants detected showed only one amino acid difference each, when compared to the major variant present in plasma. These replacements occurred at various positions within HVR1 and, except for the one present in two consecutive plasma samples of patient 2, the amino acid substitutions reverted over time. The overall d_N/d_S ratios for HVR1 were 0.030 for patient 1 and 0.323 for patient 2, strongly suggesting negative (purifying) selection

operating on this region. Consequently, a high degree of HVR1 conservation at the amino acid level was demonstrated during the course of the infection.

Selection usually occurs on some amino acid positions, not over an entire region as a whole. Consequently, we employed a high resolution approach to detect codons subject to positive or negative selection. As shown in Figure 1, several amino acid positions subject to negative selection within HVR1 were evidenced only in patient 1, whereas positive selection was detected on a single position in patient 2. Moreover, the majority of amino acid positions in HVR1 which remained unchanged during the prolonged follow-up in both patients were encoded by invariable codons, as shown by the asterisks in Figure 1. Thus, the evolution of HVR1 in patients 1 and 2 seems to be characterized by the invariability of most nucleotide positions, instead of being driven by a continuous selection-adaptation process.

To maintain its stability and function, the HVR1 must display some typical physicochemical features (8), which were observed in the variants of the HVR1 analysed in the present study. Moreover, the amino acid replacements that generated minor viral variants in the patients under study occurred in HVR1 positions predicted to be variable (8). Thus, the amino acid sequences of the viral variants detected in both patients show no particular physicochemical features that may favour their conservation over time. All the HVR1 variants of each patient were predicted to be antigenic (Figure 2), suggesting that the HVR1 of the virus infecting both patients should be able to induce an appropriate antibody response in these immunocompetent hosts.

It has been demonstrated that the genomic regions of HCV flanking HVR1 (AR) exhibit an evolutionary behaviour that is different to the one displayed by HVR1 (10, 16). The 5'-AR (nucleotides 1323–1490) and the 3'-AR (nucleotides 1572–1622) proved to be heterogeneous, and this nucleotide variability was observed both within

and between samples. In patient 1, intra-sample quasispecies diversity in the AR fluctuated significantly between consecutive samples, whereas samples corresponding to patient 2 showed increasing intra-sample diversity until 42 months of follow-up ($p<0.001$), with a subsequent decrease at the last time point tested ($p<0.001$, Table 2). In the inter-sample analysis, samples corresponding to patient 1 showed a continuous and significant reduction of quasispecies diversity between consecutive samples ($p<0.001$), whereas increasing diversification of the AR was observed in samples of patient 2 ($p<0.001$). Except for samples taken at months 29 and 61 of follow-up, nucleotide diversity showed no significant difference between HVR1 and AR in patient 1 (Table 2). In addition, intra-sample genetic distance was significantly higher in the AR than in the HVR1 in samples taken at 26 and 42 months of patient 2 ($p<0.001$). Together, the analysis of the AR demonstrated that the particular conservation of HVR1 in both patients was associated with some degree of variability in its flanking regions.

Finally, the AR showed a higher frequency of amino acid replacements than the HVR1 (Table 3). These were randomly distributed and exhibited no fixation over time (data not shown). The overall d_N/d_S ratios for the AR regions proved to be <1 (0.310 and 0.164 for patients 1 and 2, respectively). In fact, as shown in Figure 1, in the high resolution analysis several sites subject to purifying selection within AR were detected, along with some positively selected positions.

Discussion

In this study, we demonstrate an extraordinary intra-host conservation of HVR1 over time in two paediatric patients who became perinatally infected with HCV. Our results are striking, since HVR1 is otherwise characterized by a high degree of heterogeneity.

Numerous polymorphisms were detected in only one clone of some samples; this observation might be attributed to sporadic nucleotide substitutions introduced during the reverse transcription and/or amplification reactions, leading to sporadic non-synonymous changes. However, sporadic amino acid substitution rates were $4{,}55 \times 10^{-5}$ and $9{,}55 \times 10^{-5}$ non-synonymous substitutions/non-synonymous site/amplification cycle for samples of patients 1 and 2, respectively, that is, higher than the error rate reported for the DNA polymerase used in this study ($1{,}6 \times 10^{-6}$/nucleotide/amplification cycle), and similar to the ones calculated by other authors with the same or different regions of the HCV genome (reviewed in 5). Thus, the observed heterogeneity mirrors the underlying viral variability and does not merely represent a methodological artifact.

The continuous variable behaviour of HVR1 in the single host has been attributed to its putative role as a target for neutralizing antibodies, and therefore its conservation has been related to hypogammaglobulinemia (28) and immunodeficiency (18). In this study, the HVR1 variants detected were predicted to be antigenic and the two patients studied are immunocompetent and had no concomitant viral infection, suggesting that HVR1 conservation over time might be the result of a particular interplay between the virus and its host that guarantees HCV fitness throughout the chronic phase of infection.

Response to alpha interferon treatment has also been associated with HVR1 evolution. Indeed, Farci et al. demonstrated an association between a rapid decrease in genetic diversity of HVR1 (in amino acids) after onset of treatment and sustained therapeutic response (11). In our study, patients 1 and 2 received alpha interferon treatment for 7 and 8 months, respectively, and none of them showed a shift in HVR1 amino acid sequences, neither during treatment nor after it. Therefore, antiviral therapy does not seem to represent a selective force for HVR1 evolution in these patients either.

The analysis of the AR indicates that the virus infecting both patients displays the variability expected for any RNA virus, and further supports the hypothesis of some kind of constraint imposed to the HVR1, and to some extent to the AR, in the patients under study. Other authors have also analyzed the regions flanking HVR1, though with AR lengths different to the ones used herein. In adult patients subjected to alpha interferon treatment, Farci et al suggested that the AR evolve under purifying selection, since the nucleotide substitutions outside HVR1 were mainly synonymous (11). Moreover, Alfonso et al demonstrated that the region downstream HVR1 takes part of virus evolution as a target of selection (10). Our results are in accordance with the former, since the scarce amino acid variability observed outside HVR1 was reversible and did not reflect an evolutionary process of positive selection. Of note, the amino acid substitutions observed in our samples occurred in the same locations as the ones described by Farci et. al. and Alfonso et al., which suggests some kind of constraint on AR variability, too.

The association between viral evolution and the course of chronic HCV infection in children has been also evaluated by other authors. Gerotto et al. reported a gradual diversification and an increase in HVR1 complexity in children with chronic HCV infection, which was independent of ALT values (17). In contrast, Farci et al. showed that patients with constantly elevated serum ALT values display an almost conserved HVR1 during a prolonged follow-up and associated this biochemical behaviour with an intense cellular, but not humoral, immune response (16). Our results support the latter observation, since both patients displayed high ALT values throughout follow-up, concomitant with HVR1 conservation. However, liver biopsies obtained from these two paediatric patients before treatment onset showed mild or moderate signs of hepatitis, as

we observed for other children with chronic hepatitis C who displayed high HVR1 variability (unpublished results).

Finally, the high viral loads displayed by our patients are a clear evidence of active viral replication, which is in contrast with the poor HVR1 evolution during a follow-up of nearly 5 years. Two potential explanations can account for the lack of HVR1 variability in these immunocompetent hosts, which was not even affected by antiviral therapy. On the one hand, there might be purifying selection operating on HVR1 variants after their adaptation to the host's environment. It should be kept in mind that both patients were born to HCV positive mothers, although no samples obtained before the age of 2 and 3 years, respectively were available. Thus, it could be possible that the conserved pattern of HVR1 evolution is the result of a process of viral adaptation to the host occurring during the first years of life, which could not be evaluated. This hypothesis is further supported by the variability observed in the regions flanking HVR1, along with the continuous detection of identical HVR1 sequences in the same patient. On the other hand, and given that both children became perinatally infected, it seems possible that the conservation of HVR1 arises from a phenomenon of immune tolerance developed in both patients early in life. This would suppress immune responses against circulating HVR1 variants as a selective force driving the evolution of this region, without affecting the general immune status of the host.

The GenBank/EMBL/DDBJ accession numbers of the sequences determined in this work are EU045934 to EU046001.

Acknowledgements M.V.P. and R.H.C. are members of the National Research Council (CONICET) Research Career Program, M.I.G. was supported by a fellowship from

CONICET. This study was funded in part by grants from ANPCyT (grant number PICT 25344) and CONICET (grant number PIP 5359).

References

1. Scannell KMW, Carolyn C.; Seeff, Leonard B., compilers. Management of Hepatitis C: 2002. [bibliography on the Internet]. Bethesda (MD): National Library of Medicine (US); [2002 Jun 10] . [465 screens]. (Current bibliographies in medicine; no. 2002-3). 4841 citations from January 1997 through June 2002. Available from: http://www.nlm.nih.gov/pubs/cbm/hepatitis_c_2002.html 2002.
2. Gismondi MI, Turazza EI, Grinstein S, Galoppo MC, Preciado MV. Hepatitis C virus infection in infants and children from Argentina. J Clin Microbiol 2004;42(3):1199-202.
3. Jara P, Resti M, Hierro L, Giacchino R, Barbera C, Zancan L, et al. Chronic Hepatitis C Virus Infection in Childhood: Clinical Patterns and Evolution in 224 White Children. Clinical infectious diseases 2006;36:275-280.
4. Simmonds P. Genetic diversity and evolution of hepatitis C virus – 15 years on. J Gen Virol 2004;85:3173-3188.
5. Smith D, McAllister J, Casino C, Simmonds P. Virus `quasispecies´: making a mountain out of a molehill? J Gen Virol 1997;78:1511-1519.
6. Stumpf MPH, Pybus OG. Genetic diversity and models of viral evolution for the hepatitis C virus. FEMS Microbiology Letters 2002;2002(214):143-152.
7. Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, Layden TJ, et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science 1998;282(5386):103-107.
8. Penin F, Combet C, Germanidis G, Frainais PO, Deleage G, Pawlotsky JM. Conservation of the conformation and positive charges of hepatitis C virus E2 envelope glycoprotein hypervariable region 1 points to a role in cell attachment. J Virol 2001;75(12):703-5710.
9. Hino K, Korenaga M, Orito E, Katoh Y, Yamaguchi Y, Ren F, et al. Constrained genomic and conformational variability of the hypervariable region 1 of the hepatitis C virus in chronically infected patients. J Viral Hepatitis 2002;9:194-201.
10. Alfonso V, Mbayed VA, Sookoian S, Campos RH. Intra-host evolutionary dynamics of hepatitis C virus in treated patients. J Gen Virol 2005;86:2781-2786.
11. Farci P, Strazzera R, Alter HJ, Farci S, Degioannis D, Coiana A, et al. Early changes in hepatitis C viral quasispecies during interferon therapy predict the therapeutic outcome. Proc Natl Acad Sci 2002;99(5):3081-3086.
12. Pawlotsky JM, Germanidis G, Frainais PO, Bouvier M, Soulier A, Pellerin M, et al. Evolution of the hepatitis C virus second envelope protein hypervariable region in chronically infected patients receiving alpha interferon therapy. J Virol 1999;73(8):6490-6499.
13. Chambers TJ, Fan X, Droll DA, Hembrador E, Slater T, Nickells MW, et al. Quasispecies heterogeneity within the E1/E2 region as a pretreatment variable during pegylated interferon therapy of chronic hepatitis C virus infection. J Virol 2005;79(5):3071-3083.

14. Manzin A, Solforosi L, Petrelli E, Macarri G, Tosone G, Piazza M, et al. Evolution of hypervariable region 1 of hepatitis C virus in primary infection. J Virol 1998;72(7):6271-6276.
15. Manzin A, Solforosi L, Debiaggi M, Zara F, Tanzi E, Romanò L, et al. Dominant role of host selective pressure in driving hepatitis C virus evolution in perinatal infection. J Virol 2000;74(9):4327-4334.
16. Farci P, Quint I, Farci S, Alter HJ, Strazzera R, Palomba E, et al. Evolution of hepatitis C viral quasispecies and hepatic injury in perinatally infected children followed prospectively. Proc Natl Acad Sci 2006;103(22):8475-8480.
17. Gerotto M, Resti M, Dal Pero F, Migliorato I, Alberti A, Bortolotti F. Evolution of hepatitis C virus quasispecies in children with chronic hepatitis C. Infection 2006;34(62-65).
18. Pollack H, Zhiying H, Hughes AL, Borkowsky W. Perinatal transmission and viral evolution of hepatitis C virus quasispecies in infants coinfected with HIV. J Acquir Immune Defic Syndr 2004;36:890-899.
19. Sambrook J, Russell D. Molecular Cloning: A Laboratory Manual. Third Edition ed. New York: Cold Spring Harbor Laboratory Press; 2001.
20. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 1997;24:4876-4882.
21. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics 1998;14(9):817-818.
22. Swofford DL. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). In. Version 4.0b10 ed. Sunderland, Massachusetts: Sinauer Associates; 2003.
23. Nei M, Gojobori T. Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 1986;3:418-426.
24. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 2004;5:150-163.
25. Kosakovsky Pond SL, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. Bioinformatics 2005;21:676-679.
26. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999;41:95-98.
27. Deleage G, Combet C, Blanchet C, Geourjon C. ANTHEPROT: An integrated protein sequence analysis software with client/server capabilities. Computers in biology and medicine 2001;31(4):259-267.
28. Booth JC, Kumar U, Webster D, Monjardino J, Thomas HC. Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemic patients. Hepatology 1998;27(1):223.

Table 1. Description of the samples analyzed in this study.

Patient	Sample Nº*	ALT (IU/L)	Serum HCV RNA (log_{10} IU/mL)	α-IFN treatment	Nº of sequenced clones
1	0	181	5.36	No	9
	4	354	5.52	No	
	18	194	5.66	No	
	27	215	ND	Yes	11
	29	291	5.49	Yes	13
	47	199	5.53	No	11
	61	177	5.50	No	10
2	0	66	>5.93	No	11
	10	115	ND	No	
	12	112	5.84	No	
	17	132	>5.93	No	
	19	137	5.91	No	
	26	221	5.88	Yes	11
	42	87	>5.93	No	12
	63	56	>5.93	No	6

* Sample numbering in months of follow-up.

ND: not determined

Table 2. Median intra- and intersample nucleotide diversity calculated for HVR1 (A) and its adjacent regions (B) of both patients studied.*

Patient 1

A

Sample	0 mo	27 mo	29 mo	47 mo	61 mo
0 mo	**1.25 (0.00-2.51)**				
27 mo	1.24 (0.00-2.52)	**0.00 (0.00-2.52)**			
29 mo	1.24 (0.00-2.51)	0.00 (0.00-2.51)	**0.00 (0.00-1.24)**		
47 mo	2.49 (1.24-2.51)	1.24 (1.24-2.51)	1.24 (1.24-2.49)	**0.00 (0.00-0.00)**	
61 mo	1.25 (0.00-3.83)	1.24 (0.00-3.83)	1.24 (1.24-2.49)	2.50 (1.24-3.80)	**1.24 (0.00-3.81)**

B

Sample	0 mo	27 mo	29 mo	47 mo	61 mo
0 mo	**0.92 (0.00-2.34)**				
27 mo	0.92 (0.00-2.33)	**0.46 (0.00-1.39)**			
29 mo	0.92 (0.00-2.33)	0.92 (0.00-1.86)	**0.46 (0.00-1.86)**		
47 mo	0.46 (0.00-1.86)	0.46 (0.00-1.39)	0.46 (0.00-1.39)	**0.00 (0.00-0.46)**	
61 mo	0.92 (0.00-2.34)	0.46 (0.00-1.86)	0.92 (0.00-1.86)	0.00 (0.00-1.39)	**0.46 (0.00-1.86)**

Patient 2

A

Sample	0 mo	26 mo	42 mo	63 mo
0 mo	**0.00 (0.00-2.72)**			
26 mo	0.00 (0.00-4.16)	**0.00 (0.00-4.16)**		
42 mo	0.00 (0.00-2.72)	0.00 (0.00-4.16)	**0.00 (0.00-2.72)**	
63 mo	0.00 (0.00-1.33)	0.00 (0.00-2.72)	0.00 (0.00-1.33)	**0.00 (0.00-0.00)**

B

Sample	0 mo	26 mo	42 mo	63 mo
0 mo	**0.46 (0.00-1.41)**			
26 mo	0.93 (0.00-2.41)	**1.41 (0.00-2.92)**		
42 mo	1.41 (0.00-3.97)	1.90 (0.00-4.51)	**1.90 (0.00-3.44)**	
63 mo	1.90 (1.41-2.91)	2.39 (1.41-3.43)	2.40 (1.91-3.95)	**0.00 (0.00-0.46)**

*Median genetic distance (min.–max.), expressed as nucleotide substitutions/100 sites. Boldface represents intra-sample genetic distance.

Sample identification is given according to the month of follow-up in which the sample was obtained.

Table 3. Variable amino acid positions within the fragment analyzed.

Patient	Amino acid position[‡]																								
	329	**330**	**332**	**335**	**340**	**345**	**348**	**351**	**354**	**357**	**367**	**368**	**372**	**377**	**379**	**391**	**395**	**397**	**402**	**411**	**420**	**421**	**423**	**424**
1	T	A	L	A	I	L	I	A	G	A	N	W	L	L	A	T	A	A	I	I	W	H	N	S
		V		T	N		V	T		T			P	F		S	T				*[§]			N
2	T	A	L	A	I	L	I	A	G	A	N	W	L	L	A	S	A	A	F	I	W	H	N	S
	A		S			V			*		S	R			S		V	V	S	V	R	R	D	
	I										D	*									*			

[‡] Numbering of the amino acid position corresponds to HCV-1 prototype sequence (GenBank accession number M62321). For each patient, the first line represents the most abundant amino acid at each position. Gray shading denotes variable amino acid positions within HVR1 region.

[§] An asterisk indicates end of translation.

Figure 1 Schematic representation of selection operating at the codon level, for patient 1 (**A**) and patient 2 (**B**) . The graph bars represent evidence of positive selection (+), negative selection (-) or no evidence of selection (no bar), over the entire fragment analyzed. The shaded area delimits HVR1. Asterisks denote amino acid positions encoded by invariable codons within HVR1.

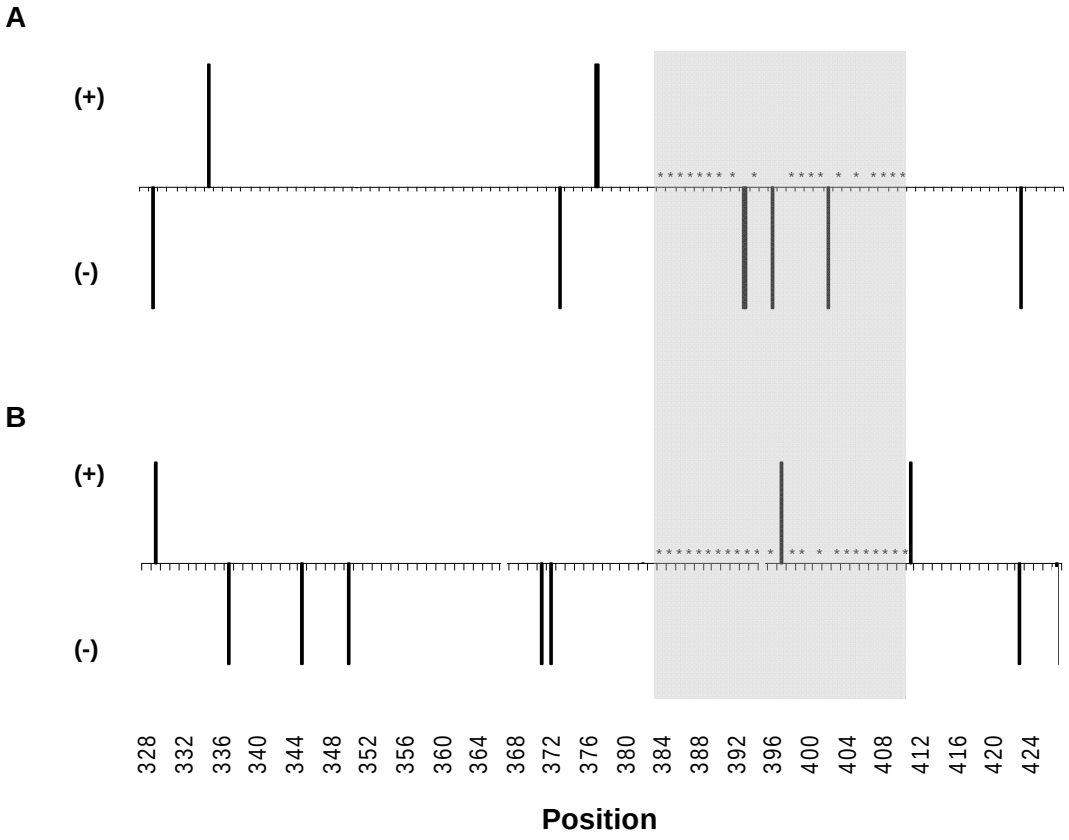

Figure 1 Predicted antigenicity profiles of the HVR1 variants detected in patient 1 (**A**) and patient 2 (**B**) during follow-up. The ruler indicates amino acid position within HVR1.