

# Genetic evolution of structural region of hepatitis C virus in primary infection

Song Chen, Yu-Ming Wang

**Song Chen, Yu-Ming Wang**, Department of Infectious Diseases, Southwest Hospital, Third Military Medical University, Chongqing 400038, China

**Supported by** the Natural Science Foundation of China, No. 3987069

**Correspondence to:** Song Chen, Department of Infectious Diseases, Southwest Hospital, Third Military Medical University, 30 Gaotanyan Zhengjie, Shapingba District, Chongqing 400038, China. a65427757@hotmail.com

**Telephone:** +86-23-687754141

**Received** 2001-09-26 **Accepted** 2001-11-27

## Abstract

**AIM:** To investigate the dynamics of hepatitis C virus (HCV) variability through putative envelope genes during primary infection and the mechanism of viral genetic evolution in infected hosts.

**METHODS:** Serial serum samples prospectively collected for 12 to 34 months from a cohort of acutely HCV-infected individuals were obtained, and a 1-kb fragment spanning E1 and the 5' half of E2, including Thirty-three cloned cDNAs representing each specimen were assessed by a method that combined a single-stranded conformational polymorphism (SSCP) and heteroduplex analysis (HDA) method to determine the number of clonotypes hypervariable region, was amplified by reverse transcriptase PCR and cloned. Nonsynonymous mutations per nonsynonymous site (dn), synonymous mutations per synonymous site (ds), dn/ds ratio and genetic distances within each sample were evaluated for intrahost evolutionary analysis.

**RESULTS:** Quasispecies complexity and sequence diversity were lower in early samples and a further increase after seroconversion, although ds value in the envelope genes was higher than dn value during primary infection. The trend, pronounced in most of samples, toward lower ds values in the E1 than in the 5' portion of E2. Quasispecies complexity was higher and E2 dn/ds ratio was a trend toward higher value in later samples during persistent viremia. We also found individual features of HCV genetic evolution in different subjects who were infected with different HCV genotypes.

**CONCLUSION:** Mutations of actively replicating virus arise stochastically with certain functional constraints. A complexity quasispecies exerted by a combination of either neutral evolution or selective forces shows clear differences in individuals, and associated with HCV persistence.

Chen S, Wang YM. Genetic evolution of structural region of hepatitis C virus in primary infection. *World J Gastroenterol* 2002; 8(4):686-693

## INTRODUCTION

Hepatitis C virus (HCV) is an important cause of morbidity and mortality worldwide [1,2]. Infection with HCV becomes chronic in >80 % of the cases and is a major cause of liver cirrhosis [3] and hepatocellular carcinoma [4]. In fact, HCV-related end-stage liver disease is now the leading reason for liver transplantation. HCV has also been linked to extrahepatic diseases, such as cryoglobulinemia and glomerulonephritis [5].

One of the important characteristics of HCV is that its genome exhibits significant genetic heterogeneity as a result of the accumulation of mutations during viral replication. This high mutation rate, which is characteristic of RNA viruses, can be attributed to an error-prone RNA-dependent RNA polymerase that lacks proofreading activity [6]. Several different genotypes have been described [7]. Analogous to other RNA viruses, HCV circulates in an infected individual as a population of closely related, yet heterogeneous, sequences: the quasispecies [7-11]. The distribution of mutations has been reported to be uneven but the basis of this variability is unexplained [12]. A hot-spot for mutation is described within the genome encoding a portion of E2 termed the hypervariable region (HVR), which encodes neutralizing epitopes [13,14]. It is postulated that genetic variation, translated into protein variability results in the production of HCV quasispecies. Interaction between host and virus selects those quasispecies better adapted to survival.

To date, the vast majority of studies describing HCV quasispecies evolution in humans have focused on the HVR1 of the HCV genome. In contrast, there have been very few longitudinal analyses of multiple HCV genes or entire structural genes in infected humans. Previous studies tracking multiple HCV genomic regions over time have been restricted to variations in few clones from infected individuals, and correlation with viral persistence [32]. We recently developed a method combining heteroduplex analysis (HAD) and single-stranded conformational polymorphism (SSCP) analysis (referred to herein as the HDA+SSCP method) that could assess genetic complexity with high sensitivity (ability to discriminate between distinct sequences) and specificity (chance that clones detected as distinct truly represent distinct sequences) and that could provide accurate estimates of genetic diversity in a previous investigation by sampling a sufficiently large number of cloned cDNAs [15]. In the prospective study we investigated the genomic complexity of HCV by HDA+SSCP method of structural gene sequences in serial serum from HCV-infected individuals. In order to detect evidence of selection of HCV mutation by the host, we compared the frequency of mutations resulting in amino acid substitutions (nonsynonymous) with those that are "silent" at the protein level (synonymous).

## MATERIALS AND METHODS

### Patients

As a part of a prospective study of acute HCV infection, serial serum samples were obtained from individuals in the ALIVE

cohort of injection drug abusers (IDUs) in Chongqing. These samples were tested for antibodies to HCV by using the second-generation HCV enzyme immunoassay. Individuals were identified as seroconverters when a sample tested positive following at least one negative result. For all HCV seroconverters, The presence of HCV RNA was evaluated in sera collected 6 months before seroconversion, at seroconversion, and during subsequent semiannual visits (median of 7). HCV RNA were detected by a quantitative reverse transcriptase PCR (RT-PCR) assay (AMPLICOR HCV MONITOR; Roche Diagnostic Systems, Branchburg, N. J.), the liner range of which was determined to be 500 to 500 000 copies per ml of serum by our and other laboratories [16,17]. Liver tests, including Alanine aminotransferase (ALT) levels in serum, performed at the first clinical examination and repeated during follow-up. Hepatitis B surface antigen, anti-HBc, anti-HBe and anti-HIV Ig M were negative in all subjects. The genome subtype was determined as described previously [18]. Clinical and virological backgrounds of the subjects studied are summarized in Table 1.

**Table 1** Molecular, biochemical, and serological characterization of the nine HCV primary infections

Subject	Sample	Age(yr)/ Sex	Duration of infection (month)	Log <sub>10</sub> [HCV RNA] <sup>a</sup>	ALT level (IU/ml) <sup>b</sup>	Result of ELISA <sup>c</sup>
A		28/F				
	A1		0	7.12	116	Neg
	A2		8	6.90	69	Pos
	A3		12	6.58	92	Pos
B		30/M				
	B1		0	7.27	110	Neg
	B2		6	6.50	35	Pos
	B3		29	5.70	37	Pos
C		33/M				
	C1		0	6.41	41	Pos
	C2		24	5.25	23	Pos
	C3		34	5.14	10	Pos

<sup>a</sup> Number of HCV RNA molecules per ml plasma. Time zero is the time where the first sample was available; other time are times after time zero. <sup>b</sup> Normal value, < 40 IU/ml. <sup>c</sup> ELISA: neg, negative; pos, positive.

## METHODS

**Envelope region amplification** HCV RNA characterization was based on examination of 33 cloned cDNAs spanning the 1 025-nucleotide (nt) region thought to encode envelope protein E1 and a segment of E2, including HVR1. Total RNA was extracted from 100 µl serum by using 500µl Trizol LS Reagent (Life Technologies, Gaithersburg, Md.) at room temperature, followed by chloroform extraction and isopropanol precipitation in the presence of 20 µg glycogen (Boehringer Mannheim, Indianapolis, Ind.). The RNA pellet was washed with 75 % (vol/vol) ethanol and then air dried briefly and redissolved in 50µl diethyl pyrocarbonate-treated water with 10 mM dithiothreitol (Promega, Madison, wis.) and 5 U of RNasin ribonuclease inhibitor (Promega). After incubation at 65°C for 5 min, 5µl purified RNA was used to generate cDNA in a 20-µl reaction mixture at 37°C for 1 h with 20 U of

Moloney murine leukemia virus RT (Promega) and first-round PCR reverse primer. The entire 20-µl cDNA synthesis reaction mixture was used for the first-round PCR in a 25-µl reaction mixture containing 0.75 U Expand HF polymerase mixture (Boehringer Mannheim, Indianapolis, Ind.), 1.5 mM MgCl<sub>2</sub>, 0.2 mM concentration of deoxynucleoside triphosphates, and 500 µM concentrations of primers. The primers were outer forward (837 to 862; 5' -GCAACAGGGAAYYTDCCGGTTGCTC-3'), outer reverse (2020 to 1998; 5' -TTCATCCASGTRCAVCCRAACCA-3'), inner forward (846 to 874; 5' -AAYYTRCCCGGTTGCTCYT TYTCTA-3'), and inner reverse (1882 to 1857; 5' -GTGAAR CARTACACYGGRCCRCANAC-3'). Degenerate bases are indicated with standard codes of the international union of Pure and applied chemistry. Nucleotide positions are numbered according to the HCV-J6 sequence. One microliter of the first-round reaction mixture was added to the second-round PCR, which had the same reagents as in the first round except for primers. Thermal-cycling conditions for the inner and outer reactions were denaturation for 120 s at 94 °C, followed by 35 cycles of 45 s at 94 °C, 45 s at 60 °C, and 120 s at 72 °C (during the last 25 cycles, the elongation time was increased by 20 s per cycle).

**Cloning of cDNA and complexity analysis of 33 cloned cDNAs by gel shift** The 1-kb HCV cDNA product was ligated into vector pT-adv and used to transform *Escherichia coli* TOP 10F' competent cells (TA Cloning kit; CLONTECH Laboratories, Inc.). Transformants were detected per manufacturer's protocol, and cloning efficiency was >90 %.

Polyacrylamide gel electrophoresis (8 %) was carried out with the addition of 15 % (W/V) urea to increase the resolution. For each subject, the gel shift patterns of 33 cloned cDNAs were examined by amplifying a 573-bp region spanning HVR1 and by using a nonradioactive method that detects distinct variants within a sample by using a combination of heteroduplex analysis (HDA) and single-stranded conformational polymorphism (SSCP) on a single gel (HDA+SSCP) [15]. Sequences obtained from the serial passage were analyzed by using a divergent variant from the acute-phase sample from each subject. A clonotype is defined as two or more cloned cDNAs that have indistinguishable patterns of electrophoretic migration by HAD+SSCP. The complexity of the quasispecies was characterized by the clonotype ratio, calculated as the number of clonotypes divided by 33, the number of cloned cDNAs examined. The clonotype ratio therefore varies from 0.03 (homogenous) to 1 (highly complex).

**Nucleotide sequencing** To examine each subject's quasispecies for signature sequences (motifs uniquely shared by a group of sequences) and for evolutions in the sensitivity of the HDA+SSCP method, a subset of cloned cDNAs was identified. For each subject, at least two cloned cDNAs were selected for sequencing based on gel shift patterns: one from the majority clonotype, another from each clonotype consisting of more than 10 % of the 33 cloned cDNAs examined, and the cloned cDNA with the largest heteroduplex gel shift. Sequences were positively determined from the M13 reverse primer and negatively from T7 promoter binding sites of plasmid clones by using a PRISM 377 DNA Sequencer (version 3.3; Applied Biosystems, Inc., Foster City, Calif.). Sequences were assembled by using the ESEE3s program, and primer sequences were removed.

**Phylogenetic analysis** DNA distance matrices were calculated by using the DNADIST program, maximum-likelihood, with a transition-to-transversion ratio of 4.25 [19], and phylogenetic trees were generated by the Neighbor-joining program with



A group:

171

GCSFSIFLLALFSLCTCPASGLEHRN<sup>1</sup>ASGLYILTND<sup>2</sup>CNSGSIVYEADDEVILHLP<sup>3</sup>GCVPCTATGNQ<sup>4</sup>TSCWTPVSP<sup>5</sup>TVAVRHPGATTASIRNHVDMLVG  
 -----T-----H-----  
 -----T-----K-----D-----S-----G-----  
 -----K-----D-----S-----G-----  
 -----Y-T-V-----V-----Y-----V-----  
 -----Y-T-V-----V-----Y-----V-----

AATLCCALYIGDL<sup>6</sup>CGAVFLVGQAFTFRPRRH<sup>7</sup>TTVQTCN<sup>8</sup>CSYIPGHISGHRMAWD<sup>9</sup>MMMNWSPAIGLLISHLMRLPQTFFDLVIGAHWGV<sup>10</sup>MAGLAYFSM  
 -----S-----  
 -----S-V-----A-----V-----R-----IT-----  
 -----S-V-----A-----T-----  
 -----S-V-----Q-----L-----  
 -----S-V-----Q-----L-----

← F1-384 E2 →

QGNWAKVCIVLIMFSGVDAGTHTTGGAAAYSTSGLASLFTQGP<sup>11</sup>KQNLHLVNSNGSWH<sup>12</sup>INSTALSCNDSLNTGFIAGL<sup>13</sup>VYHHKFNSTGCPARMSSCKP  
 -----S-----V-----  
 -----G-----S-Y-S-S-RTA-FV-SP-Q-----N-----F-Y-----S-----K-----  
 -----G-----N-Y-S-R-FV-SP-Q-----N-----F-Y-----S-----KY-----  
 -----G-----T-Y-SDA-RTAK-IT-S-Q-----N-----I-----F-Y-----  
 -----G-----T-Y-N-A-RTAK-IT-S-Q-----N-----I-----F-Y-----  
 HVR1 |-----|

511

ITAFKQGWGSLK<sup>14</sup>DVNI<sup>15</sup>SGPSEDRPYCWHYPPRPCDTVQALKVCGPVYCF (A1.1)  
 ----- (A1.2)  
 -----E-----T-M-S-----R-IE-PT----- (A2.1)  
 -----E-----T-M-S-----R-IE-PT-M----- (A2.2)  
 -----T-----S-S-N-SE----- (A3.1)  
 -----T-----S-S-N-SE----- (A3.2)

B group.

171

GCSFSIFLLALFSLCTCPASGLE<sup>16</sup>YRNTSGLYILTND<sup>17</sup>CNSKSIVYEADDVILHSPGCV<sup>18</sup>PCTATGNQ<sup>19</sup>TSCWTPVSP<sup>20</sup>TVAVRYPGATTASIRGHVDMLVG  
 -----K-----V-----  
 -----Q-----  
 -----Q-----M-----  
 -----G-----Q-----H-----M-----  
 -----Q-----M-----

AATLCSALYVGD<sup>21</sup>LCGAVFLVGQAFTFRPRRH<sup>22</sup>ATVQTCN<sup>23</sup>CSYIPGHISGHRMAWD<sup>24</sup>MMMNWSPA<sup>25</sup>VGLLVSHLMRLPQTFFDLVTGAHWGV<sup>26</sup>MAGLAYFSM  
 -----I-----  
 -----V-----I-----I-----  
 -----I-----T-----I-----I-----  
 -----I-----I-----V-----  
 -----I-----I-----V-----

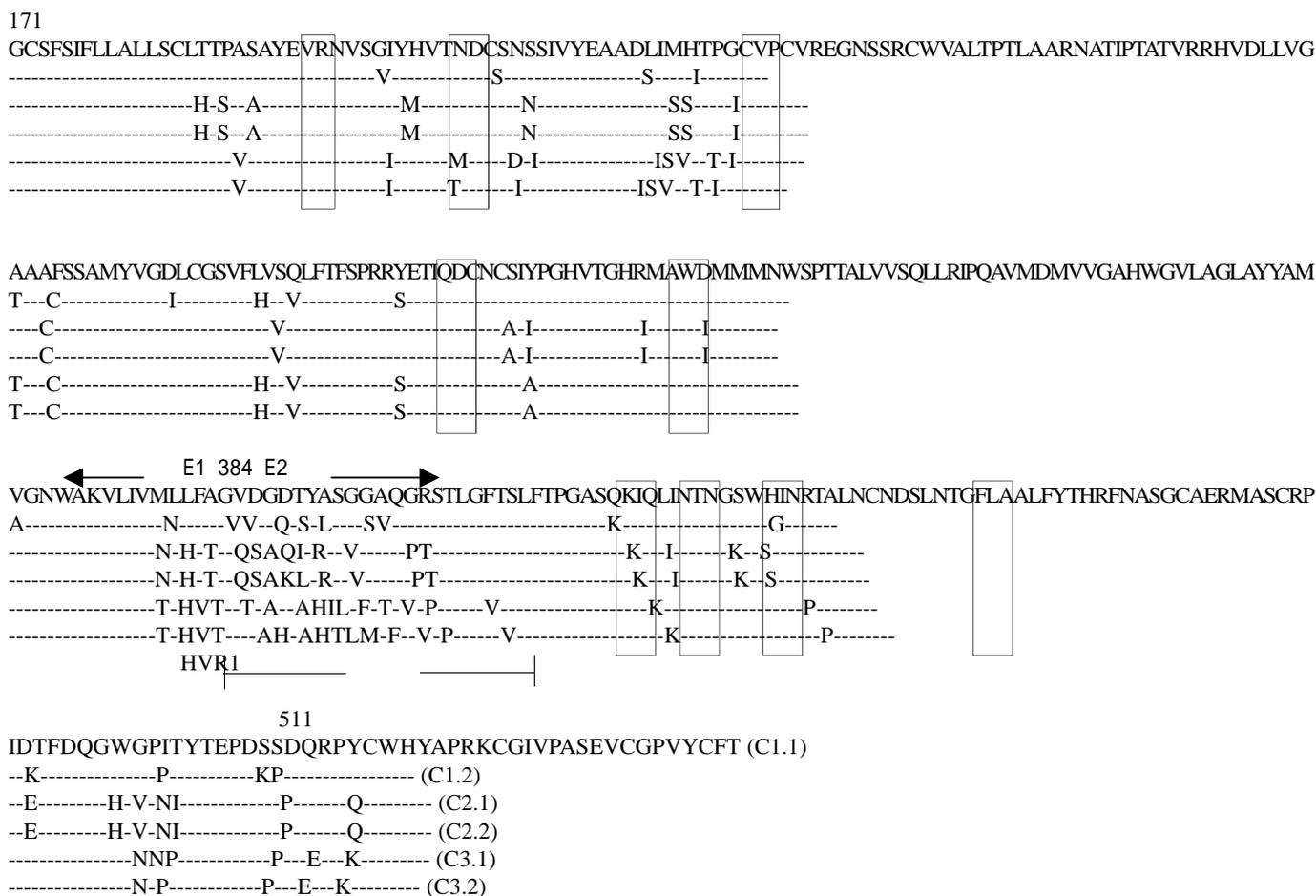
← E1-384 E2 →

QGNWAKVGI<sup>27</sup>VLIMFSGVDAGTYTTGGSAARSASGLVGLFSSGPQ<sup>28</sup>QNLHLVNSNGSWH<sup>29</sup>INSTALNCNDSLNTGFIAGL<sup>30</sup>FYH<sup>31</sup>HKFNSTGCPDRMSRCKP  
 -----R-----  
 -----T-HT-TF-R-P-S-Q-----S-----I-----SA-S-R-----  
 -----T-HT-TF-P-S-Q-----S-----I-L-SA-S-----  
 -----T-T-S-NMG-A-ITN-T-R-Q-T-----I-----I-----A-S-----  
 -----T-S-NVG-G-ITN-T-R-Q-T-----I-----I-----A-S-----  
 HVR1 |-----|

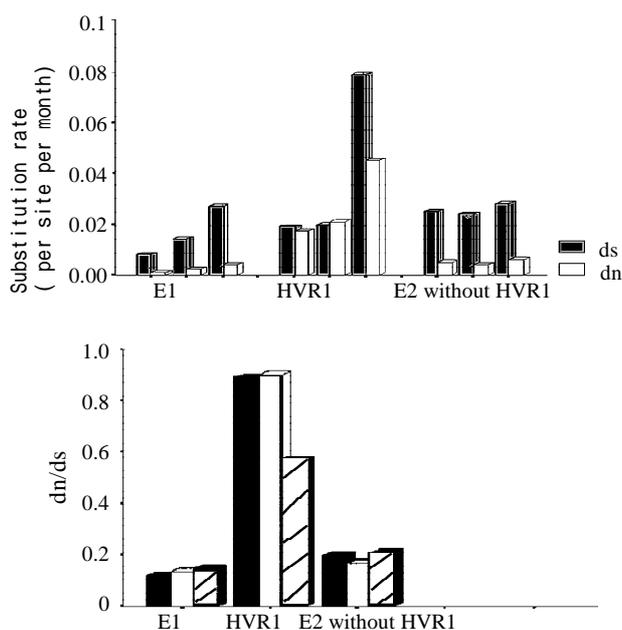
511

ITAFEQGWGSL<sup>32</sup>TDV<sup>33</sup>NVSGSSEDRPYCWHYPPRPCETVKAPTVC<sup>34</sup>GPVYCF (B1.1)  
 ----- (B1.2)  
 -----K-----K-I-P-----N-Q----- (B2.1)  
 -----K-----K-I-P-----N----- (B2.2)  
 -----D-----I-----N-Q-L----- (B3.1)  
 -----D-----N-----L----- (B3.2)

C group:



**Figure 3** Alignment of inferred amino acid sequences for the majority clonotype and each clonotype from each subjects A, B, F at different time points. In the last column, an alphabetical label is given for each subject. Period indicate identity to the amino acid at that position in the first sequence. Position of the E1 and E2 region are indicated above the alignment, whereas that of HVR1 is indicated below the alignment at the N terminus of E2. potential N-linked glycosylation site are boxed. For each subject, numbers indicate the different clones obtained.



**Figure 4** ds and dn, and dn/ds for acute-phase samples and persistent-phase samples from three individuals with HCV infection. In each panel, The individuals are presented in the order A, B.

**Detailed analysis of sequences from serial passages**

Figure 3 shows the amino acid alignment of the sequences derived from the cloned E2/HVR1 and the flanking E1 subfragment obtained from the three individuals and highlights the nonsynonymous changes. A phylogenetic study of these sequences (Figure 2) revealed that sequences from two individuals (A and B) tended to cluster tightly, segregated away from clusters of sequences from the individual C. For individuals A and B, the sequences from samples A1 to A3 or B1 to B3 maintained a low mean intrasample genetic distance: 0.0009 (A1), 0.0131 (A2), 0.0089 (A3), and 0.0014 (B1), 0.0054 (B2), 0.0046 (B3), respectively. The sequences found in the samples from individual C displayed a greater intrasample distance with late sample (C3 [collected 34 months after C1]) documenting a sharp increase (0.0214). An additional intrasample parameter evaluated in this study was the proportion of nonsynonymous substitutions per potential nonsynonymous site (dn) and synonymous substitutions per potential synonymous site (ds); It is shown in Figure 4 that ds values in the E1 region are lower than ones in E2/HVR1 at the samples from all individuals. There is the trend, observed during serial passages in three individuals, toward higher intersample dn/ds ratios in the E2/HVR1 than in E1, although

intersample dn/ds ratio were less than 1.0 with except for individual C (Figure 4). These data principally document that (i) early sequences have a low genetic divergence (oligoclonal profile), (ii) the dynamics and extent of the host selective pressure may differ among different individuals with different HCV genotypes, (iii) there are segmental effects of selective pressure on the envelope region of HCV genome.

## DISCUSSION

Sequence analysis of cDNA clones derived from PCR products from individual patients has provided important information on the genetic variability of HCV HVR1 [22-29]. However, taking into account the quasispecies nature of HCV infection and the marked heterogeneity of patients with chronic HCV infection, sequence analysis of a large number of clones would be necessary for accurate assessment of viral evolution in individuals, but because of the effort and expense, published studies obtain sequence information from a small number of colonies per subject. To overcome this potentially serious limitation, HDA and SSCP assays have been developed and proposed as means of reducing sequence analyses [30-32]. Electrophoretic analysis of SSCP has been more expedient, but its sensitivity (ability to identify distinct clones) is limited and it does not provide an estimate of genetic distance [33,34]. Heteroduplex analysis (HDA) is also convenient and provides information on both genetic complexity and distance [35-43]. However, HDA alone may not be sufficiently sensitive. Because the HVR1 PCR product is less than 200 base, the distance between heteroduplexes and homoduplex does not accurately reflect the degree of heterogeneity [30]. Our approach is a method for measuring HCV quasispecies complexity that combines HDA and SSCP in a single gel visualized with UV light. Furthermore, we repeated studies using a larger region in the exposed portion of the envelope 2 gene. Similar sensitivity to the HVR1 has been shown. The method was sensitive and specific for detecting clonotypes, and the number of clonotypes detected correlated strongly with sequence diversity when 33 cloned cDNA are assessed [15].

Artificial 'minor' quasispecies can result from *Taq* polymerase-derived errors introduced during amplification [21] or from selection during amplification and cloning procedures [44]. In this study, PCR artifacts have been minimized by the use of thermostable polymerases with proofreading function; selection phenomena are not also considered, because the cloning efficiency is very high (>90 %) and defective sequences (deletion or stop codons) are not detected. It is important that the variability of the E1 and E2/HVR1 sequences in each sample was limited to segregating polymorphisms, suggesting that the nucleotide substitutions in these sequences are spontaneous mutations.

The current study principally document that low quasispecies complexity and intrasample genetic divergence was observed in the first samples collected from these three individuals, before a specific humoral immune response was elicited, or soon after antibody production, suggests that in its early phase HCV infection may be an oligoclonal event. Despite the vulnerability of IDUs to HCV multiple transmission through a variety of unknown routes, we found on evidence of mixed HCV genotype infection, neither did we find higher quasispecies complexity or genetic distance values in the first samples from the IDU compared to the blood donor (do not shown). Sequencing of inserted E1 and E2/HVR1 clones derived from all three subjects shows that variation in sequences of the minority variants involved single-nucleotide

substitutions from majority variant, accounting for the tight clustering of sequences seen in a dendrogram. These data are consistent with quasispecies evolution from a single HCV founder strain and again point to the rarity of multiple HCV carriage in IDUs [45]. Meanwhile, We note that there were differences in quasispecies complexity and diversity in the circulation before and early after the appearance of antibody. In individual A with subtype-3b infection, there was only a single change in the E2/HVR1 region in the first sample (before seroconversion), and the variants that subsequently emerged possessed more changes suggesting that these and the dominant variant belong to one monophyletic group. By contrast, there was multiple substitution in the E2/HVR1 region in the first sample (early after seroconversion) from individual C with subtype-1b infection. Such a disparity in genetic complexity of HCV in samples at the acute stage of infection has been noted in a previous study of three patients with acute hepatitis C [23]. Data from our study suggest that difference in viral quasispecies complexity and diversity in early samples could be due to a variety of factors, including whether effective humoral responses or cellular responses have been triggered, or possibly the subtype. However, it is possible that mixed infections are in fact common but that one subtype prevails and the other becomes undetectable. It would be feasible to apply the HAD+SSCP procedure to larger numbers of serum samples to confirm the absence or paucity of multiple infection in IDUs or patient subgroup. This confirmation would imply that vaccination programs employing live, attenuated HCV vaccines might be more effective in preventing primary HCV infections than inactivated, subunit, or epitope vaccines would be.

From a mechanistic perspective, variation with the HCV genome is assumed to be caused by random mutation and selection of variants which are most fit to propagate in a given host [46,47]. Generally, synonymous changes are often thought to represent a molecular clock, independent of external pressure and expected to occur at a rate proportional to the organism's reproductive rate, Whereas nonsynonymous changes are selected by immune pressure [47]. The results shown in Figure 4 are the trend, pronounced in most of samples, toward lower ds values in the E1 than in the 5' portion of E2. Lower ds values in the 5' half of E1 than in a 3' segment of E2 have been detailed in recent studies [24,51,52]. Lower ds values may indicate that E1 region has some constraints on synonymous variation, such as RNA secondary structure or binding sites for factors that regulate replication or translation. However, In protein-coding regions, multiple forces affect the balance between fixation of silent (synonymous) mutations versus those that alter amino acid sequence (nonsynonymous). In the present study, HVR1 variation during 6-8 months of infection did not reveal a common pattern of increasing diversity within each subject, though later sequences did diverge from earlier ones, indicating the action of selective forces. In addition, there is the trend, also pronounced in a longitudinal study of three subjects each, toward higher dn/ds ratios (as a surrogate indicator for immune pressure) in the E2/HVR1 than in E1. These results suggest that selective pressure acts differently on different genomic regions. Since the high mutation rate within the hypervariable regions of HCV E2 or the V loops Of HIV gp120 could be explained by the fact that these regions are free of structural constraints, a role for the HVR1 of HCV (and other viruses) as decoy antigen has recently been proposed, stimulating a strong immune response so that a early response to a highly immunogenic region might suppress or delay the response to other less-dominant epitopes, resulting in a diversion of the immune response away from more-conserved regions [48-50], but in some cases, is ineffective for viral clearance [51].

In conclusion, HDA+SSCP assay facilitates rapid and reliable assessment of HCV quasispecies diversity. Combined with nucleotide-sequencing procedure, it provides evidence that different dynamics of quasispecies evolution is associated with viral features and its fitness for individual environments. Mutations of actively replicating virus arise stochastically with certain functional constraints. In a longitudinal study of three subjects, HVR1 variation during the first 8 months of infection did not reveal a common pattern of increasing diversity within each subject, though later sequences did diverge from earlier one. We were unable to identify an envelope sequence motif that predicts the outcome of primary infection.

## REFERENCES

- Meier V**, Mihm S, Braun PW, Ramadori G. HCV-RNA positivity in peripheral blood mononuclear cells of patients with chronic HCV infection: does it mean viral replication. *World J Gastroenterol* 2001; **7**: 228-234
- Braun PW**, Meier V, Braun F, Ramadori G. Combination of "low-dose" ribavirin and interferon alfa-2a therapy followed by interferon alfa-2a monotherapy in chronic HCV infected non-responders and relapsers after interferon alfa-2a monotherapy. *World J Gastroenterol* 2001; **7**: 222-227
- Tong MJ**, El-Fara NS, Reikes AR, Co RL. Clinical outcomes after transfusion-associated hepatitis C. *N Engl J Med* 1995; **332**: 1463-1466
- Villano SA**, Vlahov D, Nelson KE, Cohn S, Thomas DL. Persistence of viremia and the importance of long-term follow-up after acute hepatitis c infection. *Hepatology* 1999; **29**: 908-914
- Hoofnagle JH**. Hepatitis C: the clinical spectrum of disease. *Hepatology* 1997; **26**: 15s-20s
- Chung RT**, He W, Saquib A, Contreras AM, Xavier RJ, Chawla A, Wang TC, Schmidt EV. Hepatitis C virus replication is directly inhibited by IFN-alpha in a full-length binary expression system. *Proc Natl Acad Sci USA* 2001; **98**: 9847-9852
- Chen S**, Wang YM, Li CM, Fang YF. Molecular epidemiology of HCV infection in intravenous drug abusers. *Shijie Huaren Xiaohua Zazhi* 2001; **9**: 526-528
- Kato N**, Ootsuyama Y, Tanaka T, Nakagawa M, Nakazawa T, Muraio K, Ohkoshi S, Hijikata M, Shimotohno K. Marked sequence diversity in the putative envelope proteins of hepatitis C viruses. *Virus Res* 1992; **22**: 107-123
- Martell M**, Esteban JI, Quer J, Genesee J, Weiner A, Esteban R, Guardia J, Gomez J. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J Virol* 1992; **66**: 3225-3229
- Allain JP**, Dong Y, Vandamme AM, Moulton V, Salemi M. Evolutionary rate and genetic drift of hepatitis C virus are not correlated with the host immune response: studies of infected donor-recipient clusters. *J Viral* 2000; **74**: 2541-2549
- Forns X**, Purcell H, Bukh J. Quasispecies in viral persistence and pathogenesis of hepatitis C virus. *Trends Microbiol* 1999; **7**: 402-409
- Bukh J**, Miller RH, Purcell HR. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis* 1995; **15**: 41-63
- Weiner AJ**, Brauer MJ, Rosenblatt J, Richman KH, Tung J, Crawford K, Bonino F. Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology* 1991; **180**: 842-848
- Weiner AJ**, Geysen HM, Christopherson C, Hall EJ, Mason TJ, Saracco G, Bonino F. Evidence for immune selection of hepatitis C Virus (HCV) putative envelope glycoprotein variants: potential role in chronic HCV infections. *Proc Natl Acad Sci USA* 1992; **89**: 3468-3472
- Wang YM**, Ray SC, Laeyendecker O, Ticehurst RJ, Thomas D. Assessment of hepatitis C virus sequence complexity by electrophoretic mobilities of both single- and double-stranded DNAs. *J Clin Microbiol* 1998; **36**: 2982-2989
- Wang PZ**, Nie QH, Zhang ZW, Bai XG. Quantitative study of HCV RNA in different population of hepatitis C virus infection. *Shijie Huaren Xiaohua Zazhi* 2000; **8**: 1247-1250
- Park Y**, Lee KO, Oh MJ, Chai YG. Distribution of genotypes in the 5' untranslated region of hepatitis C virus in Korea. *J Med Microbiol* 1998; **47**: 643-647
- Smith DB**, Simmonds P. Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome. *J Mol Evol* 1997; **45**: 238-246
- Nei M**, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 1986; **3**: 418-426
- Smith DB**, McAllister J, Casino C and Simmonds P. Virus 'quasispecies': making a mountain out of a molehill? *J Gen Virol* 1997; **78**: 1511-1519
- 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**: 1144-1151
- Manzin A**, Solfrosi L, Petrelli E, Macarri G, Tosone G, Piazza M, Clementi M. Evolution of hypervariable region 1 of hepatitis C virus in primary infection. *J Virol* 1998; **72**: 6271-6276
- McAllister J**, Casino C, Davidson F, Power J, Lawlor E, Yap PL, Simmonds P, Smith DB. Long-term evolution of the hypervariable region of hepatitis C virus in a common-source-infected cohort. *J Virol* 1998; **72**: 4893-4905
- Ni YH**, Chang MH, Chen PJ, Lin HH, Hsu HY. Evolution of hepatitis C virus quasispecies in mothers and infants infected through mother-to-infant transmission. *J Hepatol* 1997; **26**: 967-974
- Faci P**, Shimonda A, Wang D, Cabezou T, De-Gioannis D, Strazzere A, Shimizu Y, Shapiro M, Alter HJ, Purcell RH. Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. *Proc Natl Acad Sci USA* 1996; **93**: 15394-15399
- Okamoto H**, Kojima M, Okada SI, Yoshizawa H, Iizuka H, Tanaka T, Muchmore EE, Ito Y, Mishira S. Genetic drift of hepatitis C virus during an 8.2 year infection in a chimpanzee: variability and stability. *Virology* 1992; **190**: 894-899
- Sakamoto N**, Enomoto N, Kurosaki M, Marumo F, Sato C. Sequential change of the hypervariable region of the hepatitis C virus genome in acute infection. *J Med Virol* 1994; **42**: 103-108
- Weiner AJ**, Thaler MM, Crawford K, Ching K, Kansopon J, Chien DY, Hall JE, Hu F, Houghton M. A unique predominant hepatitis C virus variant found in an infant born to a mother with multiple variants. *J Virol* 1993; **67**: 4365-4368
- Lee JH**, Stripf T, Roth WK, ZeuZem S. Non-isotopic detection of hepatitis C virus quasispecies by single strand conformation polymorphism. *J Med Virol* 1997; **53**: 245-251
- Moribe T**, Hayashi N, Kanazawa Y, Mita E, Fusamoto H, Negi M, Kaneshige T, Igimi H, Kamada T, Uchida K. Hepatitis C viral complexity detected by single-strand conformation polymorphism and response to interferon therapy. *Gastroenterology* 1995; **108**: 789-795
- Sullivan DG**, Wilson JJ, Carithers RLL, 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**: 10036-10043
- Cotton RG**. Current methods of mutation detection. *Mutat Res* 1993; **285**: 125-144
- Carrington M**, Millter T, White M, Gerrard B, Stewart C, Dean M, Mann D. Typing of HLA-DQA1 and DQB1 using

- DNA single-stranded conformational polymorphism. *Hum Immun* 1992; **33**: 208-212
- 34 **Calvo PL**, Kansopon J, Sra K, Quan S, Dinello R, Guaschino R, Calabrese G, Danielle F, Brunatto MR, Bonino F, Massaro L, Polito A, Houghton M, Weiner AJ. Hepatitis C virus heteroduplex tracking assay for genotype determination reveals diverging genotype 2 isolates in Italian hemodialysis patients. *J Clin Microbiol* 1998; **36**: 227-233
- 35 **Delwart EL**, Shpaer EG, Louwagie J, McCutchan FE, Graz M, Rubsamen WH, Mullins JJ. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science* 1993; **262**: 1257-1261
- 36 **Chen S**, Wang YM, Fang YF, Wu ZJ. Genetic complexity of the envelope region 2 (E2) of hepatitis C virus in HCV-infected individuals. *Zhonghua Weishengwuxue He Mianyixue Zazhi* 2001; **21**: 397-399
- 37 **Gavier B**, Martinez-Gonzalez MA, Riezu-Boj JJ, Lasarte JJ, Garcia N, Civeira MP, Prieto J. Viremia after one month of interferon therapy predicts treatment outcome in patients with chronic hepatitis C. *Gastroenterology* 1997; **113**: 1647-1653
- 38 **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**: 7622-7631
- 39 **Kreis S**, Whistler T. Rapid identification of measles virus strains by the heteroduplex mobility assay. *Virus Res* 1997; **47**: 197-203
- 40 **Nelson JAE**, Fiscus SA, Swanstrom R. Evolutionary variants of the human immunodeficiency virus type 1 V3 region characterized by using a heteroduplex tracking assay. *J Virol* 1997; **71**: 8750-8758
- 41 **Polyak SJ**, Faulkner G, Carithers RL, Corey L, Gretch DR. Assessment of hepatitis C virus quasispecies heterogeneity by gel shift analysis: correlation with response to interferon therapy. *J Infect Dis* 1997; **175**: 1101-1107
- 42 **Wilson JJ**, Polyak SJ, Day TD, Gretch DR. Characterization of simple and complex hepatitis C virus quasispecies by heteroduplex gel shift analysis: correlation with nucleotide sequencing. *J Gen Virol* 1995; **76**: 1763-1771
- 44 **Forns X**, Bukh J, Purcell RH, Emetsen SU. How *Escherichia coli* can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. *Proc Natl Acad Sci USA* 1997; **94**: 13909-13914
- 45 **Harris KA**, Teo CG. Diversity of hepatitis C virus quasispecies evaluated by denaturing gradient gel electrophoresis. *Clin Diagn Lab Immunol* 2001; **8**: 62-73
- 46 **Kato N**, Ootsuyama Y, Tanaka T, Nakagawa M, Nakazawa T, Muraio K, Ohkoshi S, Hijikata M, Shimotohno K. Humoral immune response to hypervariable region 1 of the putative envelope glycoprotein (gp70) of hepatitis C virus. *J Virol* 1993; **67**: 3923-3930
- 47 **Gojobori T**, Morijama EN, Kimura, M. Molecular clock of viral evolution, and the neutral theory. *Proc Natl Acad Sci USA* 1990; **87**: 10015-10018
- 48 **Zhang L**, Diaz RS, Ho DD, Mosley JW, Busch MP, Mayer A. Host-specific driving force in human immunodeficiency virus type 1 evolution in vivo. *J Virol* 1997; **71**: 2555-2561
- 49 **Wyatt R**, Sullivan N, Thali M, Repke H, Ho D, Robinson J, Posner M, Sodroski J. Functional and immunologic characterization of human immunodeficiency virus type 1 envelope glycoproteins containing deletions of the major variable regions. *J Virol* 1993; **67**: 4557-4565
- 50 **Garrity RR**, Rimmelzwaan G, Minassian A, Tsai WP, Lin G, de-Jong JJ, Goudsmit J, Nara PL. Refocusing neutralizing antibody response by targeted dampening of an immunodominant epitope. *J Immunol* 1997; **159**: 279-289
- 51 **Ray SC**, Wang YM, Laeyendecker O, Ticehurst JR, Villano SA, Thomas DL. Acute hepatitis C virus structural gene sequences as predictors of persistent viremia: hypervariable region 1 as a decoy. *J Virol* 1999; **73**: 2938-2946
- 52 **Smith DB**, Simmonds P. Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome. *J Mol Evol* 1997; **45**: 234-246

Edited by Ma JY