Increasing genetic diversity of hepatitis C virus in haemophiliacs with human immunodeficiency virus coinfection

Yasuhito Tanaka,1 Kousuke Hanada,2 Hideji Hanabusa,3 Fuat Kurbanov,1 Takashi Gojobori2 and Masashi Mizokami1

Correspondence
Yasuhito Tanaka
ytanaka@med.nagoya-cu.ac.jp

1Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Kawasumi, Mizuho, Nagoya 467-8601, Japan
2National Institute of Genetics, Yata 1111, Mishima, Shizuoka, Japan
3Ogikubo hospital, Tokyo, Japan

Patients with inherited bleeding disorders who received clotting factor concentrates before 1987 have high rates of hepatitis C virus (HCV) or HCV/human immunodeficiency virus (HIV) infection. To determine whether the persistent nature of HIV affects the genetic diversity of HCV by less selective pressure through the immunosuppression of HIV/HCV-coinfected patients, both the change of genetic diversity and selective pressure were examined in the HCV envelope genes (E1 and E2) of 325 genotype 1a subclones from eight HIV-positive and five HIV-negative patients with two time points (more than 6 years apart). To infer the genetic diversity of HCV in each patient, we used two approaches. One method was to estimate the difference of total evolutionary distances in the phylogenetic tree between the two time points, and another was to estimate the changes of genetic diversity along the time based on the coalescence theory. The two results indicate that the HIV-positive group has significantly more diverse population structure than the HIV-negative group. A comparative analysis of the synonymous and non-synonymous substitutions found that the HIV-positive group was subject to less selective pressure than the HIV-negative group. In conclusion, HIV-positive patients would have a more diversified HCV population than HIV-negative patients due to less selective pressure from the immune system.

INTRODUCTION
Increased rates of progression to end-stage liver disease, mortality and reduced treatment response rates have been well documented in haemophiliac and other groups of chronic hepatitis C virus (HCV) carriers with human immunodeficiency virus (HIV) coinfection (Bica et al., 2001; Goedert et al., 2002; Braitstein et al., 2004). Although the mechanism of liver disease progression in HIV-infected patients remains unclear, one of the important roles is assigned to immunosuppression (Goedert et al., 2002).

The estimated HCV virion half-life time was, on average, 2.7 h with pre-treatment production and clearance of $10^{12}$ virions per day (Neumann et al., 1998). Such a high rate of HCV replication, combined with lack of an error correction mechanism, results in the development of genetically diverse clones in a patient. The genetic diversity of HCV has provided critical insights into short-term outcomes, including early spontaneous viral clearance (Farci et al., 2000), interferon-associated viral clearance (Farci et al., 2002; Pawlotsky et al., 1999) and HCV emergence following liver transplantation (Lyra et al., 2002). To infer the genetic diversity of HCV in a patient, we applied two approaches. One method simply assumed that the genetic diversity of HCV is of different divergence of synonymous distance between two time points. The other method applied to coalescent analysis of genetic diversity along the time, assuming that the genetic diversity of HCV represents a heterogeneous viral population in a given carrier. To evaluate the influence on the HCV evolution exerted by the immunosuppression during persistent HIV coinfection, we examined a cohort of HCV carriers, comparing the HIV-positive and -negative groups.

Determination of the antigen-recognition regions associated with HCV-specific immune positive selection is important for understanding selective pressures underlying the evolution of HCV as well as putative therapeutic targets. In this study, we evaluated the genetic diversity of HCV and determined genomic regions associated with positive selection by comparative analysis of selective forces between HIV-positive and -negative groups in a cohort of haemophilia patients followed for more than 6 years.
METHODS

Selection of patients. The patients enrolled in the present study were a subset of a well-characterized cohort of 166 patients with haemophilia who had received non-heated plasma-derived coagulation products before 1987 and had been regularly observed since 1995 at Ogikubo hospital (Tokyo, Japan). Plasma samples from patients with known HCV and HIV serological status were stored at −80 °C. Of these patients, 57 were positive and 109 were negative for anti-HIV. After exclusion of HCV-RNA-negative and interferon-treated patients, and those with a mixture or shift in HCV subtypes during the follow-up, 13 HCV-1a-RNA-positive patients (eight HIV-positive and five HIV-negative patients) were selected at random for this study. The study protocol conformed to the 1975 Declaration of Helsinki and was approved by the Ethics Committees of each institution. Every patient gave written informed consent to participate in the virological research.

Laboratory tests. Laboratory evaluation included complete blood cell count and serum transaminases [alanine aminotransferase (ALT)], CD4+ cell counts were examined by fluorescence-activated cell sorting at SRL Inc. (Tokyo, Japan). Serum HCV-RNA levels and HIV-RNA levels were measured by a commercial PCR assay (AmpliC PCR Monitor and AmpliC PCR HCV monitor; Roche Diagnostics). The detection limits of PCR for HCV-RNA and HIV-RNA were 500 IU ml−1 [0.5 kilo international unit (KIU) ml−1] and 50 copies ml−1, respectively.

HCV-RNA isolation and amplification from the core, E1 and E2 regions. Nucleic acids were extracted from serum samples using a SepaGene RV-R Nucleic Acid Extracting kit (Sanko Junyaku) in accordance with the manufacturer’s protocol. Viral RNA was reverse-transcribed to cDNA using SuperScript II RNase H− Reverse Transcriptase (Invitrogen) and random hexamer primer (Takara Shuzo) as described previously (Ohno et al., 1997).

Partial core, E1 and E2 fragments were amplified by using PCR with primers as described previously (Tanaka et al., 2002). To reduce the number of artificial substitutions arising during PCR, Platinum Pfx DNA Polymerase (Invitrogen) with a very high fidelity was used.

Cloning and sequencing of cDNA. The amplified products were ligated into pCR-Blunt II-TOPO Vector and used to transform DH5α high-efficiency competent cells according to the manufacturer’s protocol (Invitrogen). The plasmid DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen) and the presence of the inserts confirmed by digestion with EcoRI. Sequencing was performed on more than 10 clones per patient at the baseline (1995–1997) and the end point (2002–2003). All clones were sequenced with Prism Big Dye (Applied Biosystems) in an ABI 3100 DNA automated sequencer.

Construction of phylogenetic trees. Nucleotide sequences of HCV were aligned by using the program CLUSTAL_X and molecular evolutionary analyses were conducted using Molecular Evolutionary Genetic Analysis software (MEGA version 3.0; Kumar et al., 2001). The MEGA algorithms were used to calculate the mean Tamura–Nei pairwise distance for all clones as well as a matrix of Tamura–Nei pairwise distances for each patient. To confirm the reliability of the phylogenetic tree, bootstrap resampling tests were performed 1000 times.

Genetic diversity of HCV over a time course. Two approaches were used to infer the genetic diversity of HCV in each patient. In the first approach, total evolutionary distances among a heterogeneous viral population were compared between the baseline and end point for each patient in the phylogenetic tree. The phylogenetic tree of genetic diversity was constructed by using the maximum-likelihood (ML) method and the ancestral sequence was inferred at every node using the ML method (Yang et al., 1995). As the evolutionary distance in each branch, the number of synonymous substitutions per synonymous site (synonymous distance) was estimated by the modified Nei–Gojobori method. Total synonymous distances were assumed to represent the genetic diversity of a heterogeneous viral population in each patient.

The other approach is the coalescence theory based on estimation of the genetic diversity. A consensus sequence based on the sequences of all HCV clones isolated from each patient was used as an outgroup to locate the position of the root in each phylogenetic tree. The topology of the phylogenetic tree was estimated by the neighbour-joining method (PHYLIP). Based on the topology, we constructed the phylogenetic tree and inferred the evolutionary rate by the ML method under the premise of the molecular clock (TipDate) (Rambaut, 2000). Based on the trees and the evolutionary rates estimated by TipDate, the coalescent analysis of genetic diversity was conducted for each patient using the Genie v3.5 software (Pybus et al., 2001; Pybus & Rambaut, 2002). In brief, time t was transformed to year using the HCV molecular evolutionary rate, assuming the sample-collection time to be the present. Function N(t) (effective numbers of HCV infections through time) was estimated by the ML method to infer the genetic diversity of HCV (Pybus & Rambaut, 2002). Although there are several models to infer N(t), the best-fit model was different among patients. Therefore, we chose a simplified model in which the genetic diversity was assumed to be exponentially increased over time (expansion model).

Identification of positively selective regions. Positively selected regions were identified using the modified method of Suzuki & Gojobori (2001). In brief, a phylogenetic tree of sequences from HCV clones was reconstructed in each patient by the ML method. The ancestral sequence was inferred at every node in the phylogenetic tree using the ML method (Yang et al., 1995). Then, synonymous and non-synonymous substitutions throughout the phylogenetic tree were estimated in each branch for each codon site. Here, to see the differences in selective pressure for HCV between the HIV-positive and -negative groups, we independently summed the total numbers of synonymous (Ns) and non-synonymous (Nn) substitutions occurring at each codon site of the HCV clones from either eight patients infected with HIV or five patients without HIV infection. The mean numbers of synonymous (Ns) and non-synonymous (Nn) sites were calculated for each codon site by the modified Nei–Gojobori method. The genetic distance of synonymous (dS) and non-synonymous (dN) was calculated as Ns/Cs and Nn/Cn, respectively. Although the ratio dN/dS is usually used for estimating selective pressure, we used (dS + 0.5)/(dN + 0.5) ratio instead in the present study, because no synonymous substitution was found in several codon sites. The ratio was calculated along with the sequence by the sliding-window analysis. Each window size consisted of three codons.

RESULTS

Comparison of clinical characteristics between HCV patients with and without HIV infection

When we compared clinical data between HCV patients with HIV (HIV-positive group) and without HIV (HIV-negative group), there were no significant differences of mean age, sex, putative duration of HCV infection or mean peak ALT levels (116 vs 146) (Table 1). Changes of ALT levels also were not different between these two groups. Mean peaks of HCV-RNA levels in the HIV-positive group (2300 ± 668 KIU ml−1), however, were significantly higher than those in the HIV-negative group (936 ± 423, P=}
Table 1. Clinical characteristics among HCV patients in this study

All patients are male. HCV genotype of all patients is 1a. LC, Liver cirrhosis; ALT, alanine aminotransferase; NT, not tested; +, positive; −, negative.

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Putative date of HCV infection</th>
<th>LC</th>
<th>HIV</th>
<th>AIDS</th>
<th>HIV-RNA (copies ml⁻¹)</th>
<th>CD4 at baseline (µl⁻¹)</th>
<th>HCV-RNA (KIU ml⁻¹)</th>
<th>ALT (U l⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT211</td>
<td>29</td>
<td>1982</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>130 000</td>
<td>20</td>
<td>130–2100</td>
<td>2100 yes</td>
</tr>
<tr>
<td>GM248</td>
<td>39</td>
<td>1986</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>23 000</td>
<td>110</td>
<td>290–1200</td>
<td>1200 yes</td>
</tr>
<tr>
<td>OT33</td>
<td>34</td>
<td>1982</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>2300</td>
<td>286</td>
<td>170–2000</td>
<td>2000 yes</td>
</tr>
<tr>
<td>TA92</td>
<td>32</td>
<td>1984</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>98 000</td>
<td>271</td>
<td>1300–2900</td>
<td>2900 yes</td>
</tr>
<tr>
<td>NK112</td>
<td>28</td>
<td>1982</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>100 000</td>
<td>27</td>
<td>2300–2600</td>
<td>2600 yes</td>
</tr>
<tr>
<td>YY321</td>
<td>27</td>
<td>1987</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>95 000</td>
<td>35</td>
<td>1200–1700</td>
<td>1700 no</td>
</tr>
<tr>
<td>KK202</td>
<td>19</td>
<td>1987</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>NT</td>
<td>NT</td>
<td>310–1600</td>
<td>1600 yes</td>
</tr>
<tr>
<td>KN201</td>
<td>45</td>
<td>1982</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>NT</td>
<td>NT</td>
<td>300–710</td>
<td>710 no</td>
</tr>
<tr>
<td>TS246</td>
<td>20</td>
<td>1984</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>NT</td>
<td>NT</td>
<td>230–1000</td>
<td>1000 yes</td>
</tr>
<tr>
<td>SH265</td>
<td>20</td>
<td>1985</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>NT</td>
<td>NT</td>
<td>340–470</td>
<td>470 no</td>
</tr>
<tr>
<td>ST251</td>
<td>26</td>
<td>1984</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>NT</td>
<td>NT</td>
<td>590–900</td>
<td>900 no</td>
</tr>
</tbody>
</table>

0.0019), which is consistent with previous reports (Eyster et al., 1994). Additionally, seven of eight patients in the HIV group had HCV-RNA elevation more than twice during follow-up, whereas only two patients in the HIV-negative group had HCV-RNA elevation.

For the eight HCV patients in the HIV-positive group, HIV-RNA and CD4 are shown in Table 1. Three of the patients had already developed AIDS and had very low CD4 counts (20, 27, 35 µl⁻¹), and the remaining five patients with HIV also had relatively low CD4 levels (110–286) at the baseline (1995–1997) before initiating highly active anti-retroviral therapy (HAART). Thereby, all HIV-infected patients studied were considered to be in an immunity-suppressed condition. Four patients with a CD4 count less than 200, including the three AIDS patients, received anti-HIV treatments.

Long-term intra-host diversity of HCV evaluated on distinct genomic regions

It has been shown previously that the genetic diversity of HCV changes in an oscillatory manner during the natural course of the infection (Devereux et al., 1997). Taking into account that the genetic diversity of HCV analysed at a single time point might not accurately reflect the dynamic profile of the population over time, we have examined 26 serum samples collected from 13 patients at two distinct time points with intervals of at least 6 years (6–8 years). At least 11 HCV clones were isolated from a single patient at the baseline (1995–1997) and at the end point (2002–2003) of the follow-up. Overall, 325 HCV clones were thus isolated and analysed. All of them belonged to genotype 1a. Phylogenetic relation of the HCV clones isolated from all patients is shown in Fig. 1. Assuming that HCV is composed of a heterogeneous viral population, which is evolving throughout time in a given host (carrier), we aimed to estimate the size and heterogeneity of the population. Two different methods were used to attain this aim.

First, we directly compared the genetic diversity of a heterogeneous viral population between the two time points. To do that, we estimated totals (for all patients in each of the two groups) of intra-host synonymous distances at each of the two time points. This estimation was done independently in both E1 and E2 genomic regions (Fig. 2). The increased difference from the baseline to the end point between the HIV-positive and HIV-negative groups was tested by the regression analysis, and the genetic diversity of the HIV-positive group is significantly higher than that of the HIV-negative group (P=0.043).

Second, the coalescent analysis of genetic diversity of HCV was conducted for each patient. Further, mean curves of the effective numbers of HCV infections were compared between HIV-positive and -negative groups (Fig. 3). Although the estimated mean number was initially relatively lower in the HIV-positive group, the rapid change to exponential growth, was observed several years after HIV infection in this group, whereas in the HIV-negative group, the effective number was gradually increasing throughout the period of time. The difference of exponential growth is significant (P=0.04). Hence, the result obtained by either method indicated the HIV-positive group to have higher genetic diversity of the heterogeneous viral population than the HIV-negative group, suggesting that this was due to the exposure of HIV infection.

Putative positively selective regions in the E1 and E2 regions

Since the higher genetic diversity of HCV was observed in HIV-positive patients, we further examined genetic evi-
Selective immune pressure was estimated in each, E1 and E2, gene. Some differences were observed between the HIV-positive and -negative groups (Table 2, Fig. 4). Immune epitopes (11 aa segments in the E1 and 5 aa segments in the E2 region) that were observed only in the HIV-negative group might have relatively weak antigenicity. Some of the segments were previously recognized as HCV-specific potential immunogenic targets such as cytotoxic T lymphocyte (CTL) epitopes (URL: http://hcv.lanl.gov/content/immuno/tables/ctl_summary.html), indicating that the positively selected segments estimated in the present study are associated with the immune response. On the other hand, positively selected segments around the hypervariable region (HVR1) regardless of HIV infection should have strong antigenic epitopes, suggesting little influence of the HIV coinfection on the natural immune selection targeting this region.

**DISCUSSION**

A previous meta-analysis showed a significantly elevated relative risk of severe liver disease in patients coinfected...
with HIV and HCV (Braitstein et al., 2004; Graham et al., 2001). Another cohort study of HIV/HCV-coinfected patients also indicated association of low CD4 cell count, alcohol consumption rate and age at HIV/HCV coinfection with acceleration of the liver fibrosis (Benhamou et al., 1999). The increased HCV replication in the HIV/HCV-coinfected patients would induce an intermediate immune response that is large enough to induce hepatic cell destruction and fibrosis but not enough to eradicate the virus from its reservoirs (Lai et al., 2003; Poynard et al., 2003).

In the present study, the genetic diversity of HCV was higher in the HIV-positive group compared with the HIV-negative group, which could be associated with either or both higher rate of HCV replication and HIV-associated immunosuppression, leading to less selective pressure on HCV in the HIV-positive group. The increased HCV replication by immune dysregulation in the HIV/HCV-coinfected patients might damage liver cells through apoptosis and other means such as cytokine disruption as reported previously (Puoti et al., 2001). Although several

<table>
<thead>
<tr>
<th>Region</th>
<th>PS value*</th>
<th>HCV with HIV</th>
<th>HCV without HIV</th>
</tr>
</thead>
</table>

*PS (positive selection) = (dN + 0.5)/(dS + 0.5).

**Table 2.** Putative positively selective regions

Unknown epitopes identified only in the HIV-negative group are shown in bold.
studies showed that HCV diversity decreases with the degree of HIV-related immunosuppression (Babik & Holodniy, 2003; Mao et al., 2001; Martell et al., 1992; Qin et al., 2005; Roque-Afonso et al., 2002; Toyoda et al., 1997), this might be associated with the different parameters assessed for the genetic diversity; the conclusions in most previous studies were made by analysing total numbers of HCV clones and overall genetic distances at the amino acid level.

To assess the discrepancy with our results by coalescence-based estimation, we further examined the role of HCV-targeted immune pressure. The HCV nucleotide substitution pattern was compared between subjects with intact immune system versus those with HIV-associated immunosuppression using a recently designed approach. Previous reports had used the mean of pairwise synonymous and non-synonymous distances within isolates (Ray et al., 2000; Blackard et al., 2004). Since most isolates do not have an independent evolutionary process, the mean may not represent overall genetic diversity of the heterogeneous viral population. To examine the non-redundant evolutionary process, recent methods have inferred the evolutionary process throughout the phylogenetic tree (Sheridan et al., 2004; Suzuki & Gojobori, 2001; Hanada et al., 2006). In the present study, we have applied a new approach to examine selection of HCV affected by HIV. The approach supported the theory that the diversified population of HIV-positive patients is due to less selective pressure and allows identification of specific regions indicating the presence of positive selection in HIV-negative patients compared with HIV-positive patients with immunosuppression. Although these positively selective segments, which were observed only in the HIV-negative group, might have relatively weak antigenicity, most of them were located inside potential immunogenic targets and others might be somewhat new antigen-recognition regions associated with HCV-specific immune responses. Interestingly, no influence by HIV coinfection was observed in the HVR1, which contains sequence-specific immunological B-cell epitopes that induce the production of antibodies restricted to the specific viral isolate (Kato et al., 1993), indicating that the positively selected segments regardless of HIV infection should have strong antigenic epitopes. Taken together, our findings indicate that defenseless HCV clones that are extinct in usual conditions can survive in HIV-positive patients because of less immune pressure leading to HIV infection. Consequently, the genetic diversity of HCV will be greater in HIV-positive patients. In fact, the diversity of both synonymous and non-synonymous substitutions was larger in HIV-positive patients than HIV-negative patients (data not shown).

One limitation of the present study is the general lack of functional immunological data. In this study, HCV-specific CD4 ELISPOT responses were not detected in all subjects and CD8 cell counts were not measured. Therefore, we chose to use HIV load and CD4 cell counts as surrogate markers of immune suppression. Further investigations of HCV diversity in conjunction with HCV-specific cellular responses will be required when more-sensitive immunological assays are available. Another potential limitation of the study is that only two time points were sampled for each individual. However, as we inferred the evolutionary process based on the phylogenetic tree constructed using the number of clones that were isolated through the long period of follow-up (more than 6 years), we believe that the genetic diversity of the heterogeneous viral population may represent an actual evolutionary process.

In conclusion, HIV-positive patients have more diversified HCV populations than HIV-negative patients, possibly because of reduction of selective pressure from the immune system. The positively selective regions determined in this study might be antigen-recognition regions associated with HCV-specific immune responses.

ACKNOWLEDGEMENTS

Financial support: the work was supported by a grant-in-aid from the Ministry of Health, Labour and Welfare of Japan and from the Ministry of Education, Culture, Science, Sports of Japan and Japan Science and Technology Agency. Potential conflicts of interest: we do not have any commercial or other association that may pose a conflict of interest.

REFERENCES


