VIROLOGY

Distribution of genotypes in the 5′ untranslated region of hepatitis C virus in Korea

Y. S. PARK*,†, K. O. LEE*, M. J. OH* and Y. G. CHAI†

*Department of Genetic Analysis, Seoul Clinical Laboratories (SCL), Seoul Medical Science Institute, Seoul 140-230 and †Department of Biochemistry and Molecular Biology, Hanyang University, Ansan 425-791, Korea

Hepatitis C virus (HCV) is an important human pathogen that can cause acute and chronic hepatitis, liver cirrhosis and hepatocellular carcinoma. Recently, partial and entire sequence data from HCV isolates have been reported, suggesting various genotypes of HCV. The genotype may be correlated with the progression of hepatitis and maybe a prognostic marker of treatment. Thus, the availability of an assay for typing HCV RNA is important. This study developed a convenient method for genotyping HCV into six groups by PCR-RFLP with four restriction endonucleases (BstUI, HaeIII, NcoI, RsaI) in the 5′ untranslated region (UTR) of HCV. The HCV genotypes from 169 patients with HCV infections in Korea were analysed. Two genotypes, type 1b and type 2a, accounted for 47.3% and 42.6% of HCV infections, respectively.

Introduction

The major aetiological agent of post-transfusion and community-acquired non-A, non-B hepatitis has been identified as hepatitis C virus (HCV), which is distantly related to the pestiviruses and flaviviruses [1–3]. A high rate of chronic hepatitis, including hepatocellular carcinoma, is associated with HCV infection [4]. HCV contains a positive strand RNA genome composed of c. 9500 nucleotides that can encode a large viral polyprotein of c. 3010 amino acids [5, 6].

Comparative sequence analysis of HCV genomes has shown that HCV can be grouped into distinct but related genomes from the complete genomic sequence of fragments amplified by reverse transcription-polymerase chain reaction (RT-PCR) [7–9]. Genotypes with numerous subtypes have been described by the nomenclature proposed by Simmonds et al. [10–12]. They can be classified into six major groups or types provisionally designated 1–6 and divided further into genotypes or subtypes 1a, 1b, 2a, 2b, 3a, 3b and 4, etc. HCV can be typed by RT-PCR and restriction fragment length polymorphism (RFLP) analysis of the 5′ untranslated region (UTR), which has a distinct nucleotide sequence [11, 13–15]. HCV genotyping is potentially important in viral transmission studies, HCV epidemiology, and in predicting the success of interferon treatment [16–18].

This study investigated a method of HCV genotyping in the 5′ UTR with PCR-RFLP and analysed the genotypes of 169 HCV isolates from patients with HCV infection in Korea.

Materials and methods

Serum samples

One hundred and sixty-nine blood samples were obtained from HCV patients who were found to be HCV positive by RT-PCR in Korea. To prevent contamination of HCV RNA for PCR, whole blood was collected in vacuum containers and was not separated into any other tubes.

Extraction of HCV RNA and RT-PCR

RNA was extracted from 100 μl of serum by the guanidinium isothiocyanate and phenol-chloroform method with RNAzol B solution [19]. A combined RT-PCR was used to amplify HCV RNA. Five μl of RNA extract were added to a 500-μl PCR tube containing 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 9.0, Triton X-100 0.1%, 1.5 mM MgCl2), 0.2 mM of each dNTPs (dATP, dTTP, dGTP and dCTP), 1.0 U of Thermoprime plus DNA polymerase (Advanced Biotechnologies), 5 U of AMV reverse
transcriptase, 20 U of RNasin and 10 pmols of outer primers (sense: CTGTGAGGAACTATCTGCTT; anti-sense, ACTGCGAAGCAACTTATCAGG; size: 268 bp) from the 5' UTR. Reverse transcription was performed at 42°C for 45 min before denaturation at 95°C for 3 min. Thereafter, PCR was performed for 30 cycles of 1 min at 94°C, 1.5 min at 50°C and 1 min at 72°C in a 9600 thermal cycler (Perkin-Elmer). A final extension step of 72°C for 5 min was also included. Two μl of PCR product from the outer primer reaction were removed and transferred to a 250-μl thin-wall tube containing fresh reagents and 20 pmols of inner primers (sense, TTCACGCAGAAAACGTCAAG; anti-sense, TATCAGGATCCACAAGG; size: 236 bp), and subjected to a further 25 cycles of 30 s at 94°C, 45 s at 55°C and 45 s at 72°C. The PCR products were separated by agarose 2% gel electrophoresis and visualised by staining with ethidium bromide.

**RFLP experiments and HCV genotyping**

HCV genotyping was performed by the RFLP method developed in this study. Restriction patterns of each HCV genotype were analysed with the PC/GENE computer program (Intelligentic), and four restriction endonucleases were selected (BstUI, HaeIII, NciI and RsaI). Samples (5 μl) of the nested PCR product were digested with 5 U of each restriction endonuclease (New England Biolabs) in the appropriate buffer for 4 h at each incubation temperature. The digested products were separated by electrophoresis on polyacrylamide 13% gel at 100 V for 3 h. The DNA fragments were visualised by ethidium bromide staining and compared with the patterns of each HCV genotype analysed.

**Direct sequencing**

The HCV typing by PCR-RFLP was confirmed by direct sequencing. For direct autosequencing, the PCR products produced by inner primer were purified by a QIAquick PCR purification kit (Qiagen), and sequenced by the multi-colour fluorescent technique following the manufacturer's instructions in an ABI 310 genetic analyser (PE Applied Biosystems).

<table>
<thead>
<tr>
<th>Position Size</th>
<th>-280</th>
<th>-44</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

Fig. 1. Restriction map predicted from 5' UTR for HCV genotyping. The size from position -280 to -44 in the 5' UTR of HCV was 236 bp. Types 1a and 1b have specific bands with BstUI; and type 2a with HaeIII and NciI; type 2b has no restriction site with BstUI; and type 3a has a specific band with HaeIII and RsaI, and type 3b with NciI and RsaI. Restriction sites: ↓, BstUI; ↑, HaeIII; †, NciI; ‡, RsaI.
Results

RNA was extracted from serum samples of a wide range of HCV-infected individuals and HCV genotyping was performed by the RFLP method combined with RT-PCR. The size of PCR product from the inner primers was 236 bp, and sequences of each type were

Table 1. Distribution of HCV genotypes from 5' UTR in 169 patients with HCV infection in Korea

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Number (% of cases)</th>
<th>Number (% of cases)</th>
</tr>
</thead>
<tbody>
<tr>
<td>la</td>
<td>5 (3.0)</td>
<td></td>
</tr>
<tr>
<td>lb</td>
<td>80 (47.3)</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>72 (42.6)</td>
<td></td>
</tr>
<tr>
<td>2b</td>
<td>4 (2.4)</td>
<td></td>
</tr>
<tr>
<td>Mixed infection</td>
<td>2* (1.2)</td>
<td></td>
</tr>
<tr>
<td>Untypable</td>
<td>6 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>169</td>
<td></td>
</tr>
</tbody>
</table>

*HCV types of mixed infection were identified as type la/lb and 2a/2b, and were not classified separately.

Fig. 1 shows the predicted restriction map at the 5' UTR of the six genotypes analysed in this study. The restriction sites of 5' UTR sequences of each HCV type were compared and analysed. HCV-1 (accession number; M62321), HCV-H (M67463), PT-1, H77, HC-J1 (D10749), GM1, GM2 and H90 belong to a group designated as type la. Type 1b contains HCV-J (D90208), HCV-BK (M58335), HCV-JK1 (X61596), HCV-China (L02836), HCV-T (M84754), HCV-JT (D11168), HCV-J4/83 (D13558), HCV-J4/91 (D10750), HCV-L2, HCV-JT* (D11355), HCV-N (S6220) and HCV-C2 (D10934). Type 2a contains HCV-J6 (D00944), K2A and E-B9, and type 2b contains HC-J8 (D01221) and K2B. NZL1 (D17763), HCV-K3a (D28917) and Eb-1 belong to type 3a, HCV-Tr (D26556) to type 3b, GB358 (L29608) to type 4, BE96 (L29585) to type 5, and QC26 (U33431) to type 6. Each type had a specific

Fig. 2. HCV genotyping of types la, lb, 2a and 2b by RFLP analysis. PCR products of 236 bp were digested with four restriction endonucleases BstUI, HaeIII, NciI and RsaI. Fig. 1 shows the predicted restriction map at the 5' UTR of the six genotypes analysed in this study. The restriction sites of 5' UTR sequences of each HCV type were compared and analysed. HCV-1 (accession number; M62321), HCV-H (M67463), PT-1, H77, HC-J1 (D10749), GM1, GM2 and H90 belong to a group designated as type la. Type 1b contains HCV-J (D90208), HCV-BK (M58335), HCV-JK1 (X61596), HCV-China (L02836), HCV-T (M84754), HCV-JT (D11168), HCV-J4/83 (D13558), HCV-J4/91 (D10750), HCV-L2, HCV-JT* (D11355), HCV-N (S6220) and HCV-C2 (D10934). Type 2a contains HCV-J6 (D00944), K2A and E-B9, and type 2b contains HC-J8 (D01221) and K2B. NZL1 (D17763), HCV-K3a (D28917) and Eb-1 belong to type 3a, HCV-Tr (D26556) to type 3b, GB358 (L29608) to type 4, BE96 (L29585) to type 5, and QC26 (U33431) to type 6. Each type had a specific
Type 1b
cTagcctgccatggcgttagtatga\ntgtTgtGcagcctccaggA\ncccccccctccc~gagagccatagtggtctgcgg
NciI

Type 2a
cTagcctgccatggcgttagtatgag\ntgtCgtAcagcctccaggC\ncccccccctcccgggagagccatagtggtctgcgg
RsaI  HaeIII  NciI

Type 1b
aaccggtgagtacaccggaattg\ncAggaCgacCgggtcctttc\ntggatCaacccGctcAatgc\nCTggAGattt
RsaI  NciI

Type 2a
aaccggtgagtacaccggaattgc\nGGgaAgacTgggtcctttc\nntggatAaacccActcTa\ntgccCGgCCatt
RsaI  NciI  NciI  HaeIII

Type 1b
gggcgtgccccccgcAgactg\ncctagccgagtagTgttgggT\gcgaaaggcctgt
BstUI  BstUI  HaeIII

Type 2a
gggcgtgccccccgcAagaotgctagccgagtagCg\ntgggTgcgaaaggctgt
HaeIII

Fig. 3. Nucleotide sequences of PCR products from two major types, type 1b and 2a, in Korea. Amplified products were purified and sequenced with an ABI 310 genetic analyser. The DNA sequences of two samples from each type were analysed and confirmed by reading the antisense strands.

Discussion

HCV contain a positive polarity, single-stranded RNA genome with 5' and 3' UTR. The core (C), envelope 1 (E1) and envelope 2 (E2) proteins are encoded at the 5' terminus and the non-structural proteins (NS) are encoded at the 3' terminus of the single open-reading frame of the genome [20]. HCV is typical of RNA viruses in having a quasi-species nature due to relatively high mutation rates, particularly in the envelope regions of the genomes. The mutation rate is estimated to be (1.44–1.92) × 10⁻³ base substitutions/site/year [21]. The 5' UTR and NS4B region are most conserved, while the E1 and E2 regions show more variability [22–24]. Many of the assays that use type-specific primers can be predicted to either fail to differentiate genotypes or fail to amplify sequences unrecognised when the primers were originally designed. Analysis of the core region, which is generally less conserved than 5' UTR, has been used previously for genotyping [12]. Because of the low mutation rate of the 5' UTR [13, 14], patterns of nucleotide variability in the 5' UTR were analysed. Furthermore, the 5' UTR is the region of choice for detection and quantification of HCV [25, 26], and can be conveniently used for genotyping by the RFLP method. One of the advantages of using the 5' UTR for genotyping is the existence of a few identifiable polymorphisms between major genotypes that reliably associate with examples of the known genotypes. However, the method does not detect mixed infections easily. In a situation with mixed infection, either the predominant pattern alone prevails or an indeterminate pattern emerges, as the mixed digestion patterns appear together. In this study, two samples were detected as mixed infection of type 1a/1b and 2a/2b, respectively; and six samples had restriction digestion patterns that were not typable into the predicted patterns. Nakao et al. and McOmish et al. have reported RFLP-based HCV genotyping methods [14, 27]. Nakao et al. detected only two HCV-K1 and HCV-K2 types from the NS5 region, compared with six different variable genotypes in the present assay. The method described by McOmish et al. was dependent upon the electrophoresis of radioisotope-labelled DNA fragments on acrylamide for a long period [27]. By contrast, the method used in the present study is less time-
HCV genotypes may be correlated with severity of liver disease. Genotype 1b is associated with more severe hepatitis as well as a lower response to interferon than genotypes 2a and 2b [29]. Patients with high serum levels of HCV RNA have a poor response to interferon, and patients infected with genotypes 1b have higher levels of HCV RNA than those infected with other HCV genotypes [18]. Also, HCV genotypes have different geographic distributions; genotype 1a is prevalent in the UK and the USA and genotype 1b is prevalent in Japan and Taiwan [30]. The distribution of HCV among Korean patients demonstrated mainly type 1b and 2a, as is the case in Taiwan.

In conclusion, genotyping based on 5' UTR sequence analysis is possible and may complement studies on antiviral treatment and the transmission of HCV, and may also have important implications for pathogenesis, diagnosis and vaccine development.

This work was supported by a grant from Seoul Medical Science Institute of Korea (SM97003).

References

3. Miller RH, Purcell RH. Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. Proc Natl Acad Sci USA 1990; 87: 2057-2061.