Typing hepatitis C virus by polymerase chain reaction with type-specific primers: application to clinical surveys and tracing infectious sources

Hiroaki Okamoto,1 Yasushi Sugiyama,2 Shunichi Okada,3 Kiyohiko Kurai,1 Yoshihiro Akahane,3 Yoshiki Sugai,4 Takeshi Tanaka,5 Koei Sato,5 Fumio Tsuda,6 Yuzo Miyakawa7 and Makoto Mayumi1

1Immunology Division, Jichi Medical School, Tochigi-Ken 329-04, 2Department of Hygiene, University of Hiroshima, Hiroshima-Ken 734, 3First Department of Internal Medicine, Yamanashi Medical College, Yamanashi-Ken 409-38, 4Department of Internal Medicine, Iwaki Kyoritsu General Hospital, Fukushima-Ken 973, 5Japanese Red Cross Blood Center, Saitama-Ken 362, 6Department of Immunology, The Kitasato Institute, Tokyo 108 and 7Mita Institute, Tokyo 108, Japan

Based on variation in nucleotide sequence within restricted regions in the putative C (core) gene of hepatitis C virus (HCV), four groups of HCV have been postulated in a panel of 44 HCV isolates. They were provisionally designated types I, II, III and IV. A method for typing HCV was developed, depending on the amplification of a C gene sequence by polymerase chain reaction using a universal primer (sense) and a mixture of four type-specific primers (antisense). HCV types were determined by the size of the products specific to each of them. Type II was found in HCV samples from 131 (82%) of 159 blood donors, more often than in those from 48 (60%) of 80 patients with non-A, non-B (NANB) liver disease in Japan (P<0.01). In 11 haemophiliacs who had received imported coagulation factor concentrates, type I was found in five, as against type II in four. Double infection with two different HCV types was found in two patients with chronic NANB liver disease (types I and II; II and III) and two haemophiliacs (types I and II; I and III). HCV types were identical in mother and baby in each of two examples of perinatal transmission, and were also identical in donor and recipient in a case of accidental needle exposure.

Introduction

Discovery of an RNA virus (Choo et al., 1989), now accepted as hepatitis C virus (HCV), and application of a recombinant viral protein to the serological diagnosis of HCV infection (Kuo et al., 1989), represent major breakthroughs in the study of non-A, non-B (NANB) hepatitis. It is now clear that HCV accounts for most cases of acute and chronic NANB liver disease (Bruix et al., 1989; Alter et al., 1989a), and infects around 1% of the general population world-wide (Esteban et al., 1989; Stevens et al., 1990; Miyamura et al., 1990).

To date, the entire nucleotide sequence is available for at least four HCV isolates (Kato et al., 1990; Choo et al., 1991; Takamizawa et al., 1991; Okamoto et al., 1991). They are identical in only 68-1 to 91.8% of the base sequence, suggesting more than one type of HCV. Different types of HCV have been proposed on the basis of sequence variations in the entire genome (Kato et al., 1990; Choo et al., 1991; Okamoto et al., 1991) or limited regions (Kubo et al., 1989; Enomoto et al. 1990; Okamoto et al., 1990b; Takeuchi et al., 1990; Ogata et al., 1991), or of differences in the capacity of certain oligonucleotide primers in amplifying HCV sequences by polymerase chain reaction (PCR) (Kaneko et al. 1990; Farci et al., 1991). HCV types could have different geographical distributions, and this might have corresponding clinical implications.

Based on the comparison of 44 HCV isolates, with respect to a portion of the putative C (core or capsid) gene, universal and type-specific sequences were deduced in arbitrary regions 20 nucleotides (nt) in length. These were used as primers in typing HCV by PCR, a process designed to generate products of different sizes that served to distinguish four HCV types. The method was applied to surveys of different HCV types in blood donors and patients with NANB hepatitis or haemophilia in Japan. It was also used to trace the route of HCV infection in cases of perinatal transmission and accidental infection after exposure to needles.

Methods

Plasma and serum samples. An initial series of samples was obtained from 30 donors and nine patients with chronic NANB liver disease. All of them were positive for anti-C100-3 (HCV Ab ELISA Test, Ortho Diagnostic Systems) and contained HCV RNA detectable by the
double PCR method with two pairs of external and internal (nested) primers deduced from the 5'-non-coding region (Okamoto et al., 1990a). cDNA clones were obtained by PCR and sequenced for nt 126 to 506, which covered two-thirds of the putative C gene (nt 1 to 573); nt were numbered from the start of the single open reading frame of the HCV genome. These 39 samples and the five reported by others (Kato et al., 1990; Takeuchi et al., 1990; Choo et al., 1991; Ogata et al., 1991; Takamizawa et al., 1991), making a panel of 44 HCV isolates, were analysed for universal or type-specific sequences 20 nt in length.

Another series of samples was obtained from 159 blood donors, 80 patients with acute or chronic NANB liver disease and 11 haemophiliacs, all of whom were positive for anti-C100-3 and HCV RNA. Haemophiliacs were on maintenance therapy with coagulation factor concentrates, either produced in Japan or imported from the United States, at a time when heat inactivation of NANB hepatitis viruses had not yet been introduced. Most of them had been transfused repeatedly. Serum samples were obtained from mothers and babies in each of two families in which perinatal transmission of HCV infection was suspected based on the detection of anti-C100-3 and HCV RNA. Also tested were sera from a patient on maintenance haemodialysis and a staff member who contracted NANB hepatitis after accidental needle exposure to the patient's blood. Anti-C100-3 and HCV RNA were positive in sera from both the patient and the staff member. HCV in these 256 samples was typed by PCR with type-specific primers.

Extraction of HCV RNA and synthesis of cDNA. RNA was extracted from plasma or serum samples by the method described previously (Okamoto et al., 1990b), denatured at 70 °C for 1 min, and converted to cDNA with appropriate synthetic primers and cloned Moloney murine leukaemia virus reverse transcriptase (Superscript, BRL).

Universal and type-specific primers for PCR. Primers no. 23 (5'-TAGATTGGGTGTGCAGGGCGG3'; nt 126 to 145, sense) and no. 122 (5'-AGGTCCCTGTTGATTTAATT3'; nt 487 to 506, antisense) were deduced from the putative C gene of the HC-J1 isolate (Okamoto et al. 1990b). They are similar, with minor substitutions, to the other eight isolates for which corresponding sequences have been reported (Kato et al., 1990; Okamoto et al., 1990b; Takeuchi et al., 1990; Choo et al., 1991; Ogata et al., 1991; Okamoto et al., 1991; Takamizawa et al., 1991).

Primers no. 104 (5'-AGGAAGACCTTCGGACGCTTC3'; nt 148 to 167, sense) and no. 186 (5'-ATGATCCTGCCAGATGGTGC3'; nt 391 to 410, antisense) were deduced from the HC-J1 isolate and no. 256 (5'-CGCCGGACTAGGAGACTCTC3'; nt 139 to 158, sense) from the HC-J4 isolate (Okamoto et al., 1990b); they were similar in the 44 isolates in the panel with occasional, minor substitution of ≤ 2 nt as shown in Fig. 1. Antisense primers specific for the four HCV types were: no. 132 (5'-TGCTTGGGATAGGCTTC3'; nt 185 to 204, type 1), no. 158 (5'-CTCTTTGGCCTTGGTGCT3'; nt 167 to 186, type 2, sense), no. 148 (5'-CTGACTGGTGGAGTCGCG3'; nt 158 to 177, type 3), and no. 167 (5'-GAGGACCTTCGGTCGTC3'; nt 177 to 196, type 4).

Fig. 1. Nucleotide sequences and positions of universal and type-specific primers. A total of 44 HCV isolates, including nine previously reported, were compared for seven sequences of 20 nt representing various parts of the putative C gene of HCV. Three sequences conserved by all of them are shown at each end. Four sequences in the middle, related within each type, are shaded. The sequences representing them, indicated at the top, were adopted as universal (no. 256, no. 104 and no. 186) or type-specific primers (no. 132 (type 1), no. 133 (II), no. 134 (III), no. 135 (IV)). S, sense; A, antisense. The origins of the nine reported isolates are: clone no. 1, HC-J1 (Okamoto et al., 1990b); no. 5, HCV-1 (Choo et al., 1991); no. 6, H strain (Ogata et al., 1991); no. 7, HC-J4 (Okamoto et al., 1990b); no. 27, HCV-JH (Takeuchi et al., 1990); no. 28, HCV-I (Kato et al., 1990); no. 29, HCV-BK (Takamizawa et al., 1991); no. 33, HC-J6 (Okamoto et al., 1991); no. 40, HC-J7 (DDBJ, GenBank and EMBL accession number D00830).
Fig. 2. Strategy for typing HCV by PCR with type-specific primers. The genomic organization proposed for HCV is indicated at the top, and has been deduced from a postulated similarity to flaviviruses (Okamoto et al., 1990b; Takeuchi et al., 1990; Choo et al., 1991; Takamizawa et al., 1991). cDNA prepared by priming with no. 186 was amplified by the first PCR with universal primers (no. 256 and no. 186) to obtain a fragment of 272 bp as indicated by the open block. The products were then amplified by the second PCR with a universal primer (no. 104) and a mixture of four type-specific primers (no. 132 to no. 135). Shaded blocks indicate products of different nucleotide lengths that were specific to each of the four HCV types.

Amplification, cloning and sequencing of cDNA. To determine the sequence of a part of the C gene in 39 sera in the panel, HCV cDNA was primed with no. 122. It was amplified by PCR in a DNA thermal cycler (Perkin-Elmer Cetus) for 35 cycles with a GeneAmp DNA amplification reagent kit (Perkin-Elmer Cetus) using universal primers (no. 23 and no. 122), which generated a product of 381 bp representing nt 126 to 506 of the C gene. Each reaction cycle included denaturation at 94 °C for 1 min, primer annealing at 55 °C for 1.5 min and primer extension at 72 °C for 2 min. Products of PCR were treated with T4 polynucleotide kinase (Takara Biochemicals) and cloned into the HincII site of M13mp19 phage vector. cDNA clones were sequenced in both directions by the dideoxynucleotide chain termination method by the use of Sequenase version 2.0 (United States Biochemicals).

Typing HCV by PCR with type-specific primers. HCV cDNA from 256 test sera was primed with no. 186, and amplified by PCR in two stages (Fig. 2). The first PCR was performed for 35 cycles with universal primers (no. 256 and no. 186) with a reaction cycle as described above. A 1/50 amount of the products was subjected to the second PCR for 30 cycles using a universal primer (no. 104, sense) and a mixture of four type-specific primers (no. 132, no. 133, no. 134 and no. 135, antisense) with denaturation at 94 °C for 1 min, primer annealing at 60 °C for 1 min and primer extension at 72 °C for 1.5 min in each reaction cycle. The products of the second PCR were subjected to electrophoresis on a composite agarose gel made from 1.5% NuSieve and 1.5% SeaKem (FMC BioProducts), stained with ethidium bromide and observed under u.v. light.

In the second stage of PCR, type I HCV RNA was amplified only with the type I-specific primer, but not with the type II, III or IV primers. The specificity in PCR amplification of type-specific primers was confirmed also for type II, III and IV HCV RNA.

Statistical analyses. Differences in the prevalence of a certain HCV type between two groups were evaluated by Chi-squared analysis.

Results

Typing HCV by PCR with type-specific primers

A total of 44 HCV isolates, including nine previously reported (Kato et al., 1990; Okamoto et al., 1990b; Takeuchi et al., 1990; Choo et al., 1991; Ogata et al., 1991; Okamoto et al., 1991; Takamizawa et al., 1991), were compared for sequence identity in seven arbitrary regions spanning 20 nt within a part of the putative C gene [nt 126 to 506 (381 bp)]. Three regions in each of them demonstrated an infrequent substitution of ≤ 2 nt, and the remaining four regions showed appreciable differences (Fig. 1). Based on the similarity in nucleotide sequence within a limited region, the 44 isolates were classified into four types, provisionally designated I, II, III and IV. These conserved and variable regions of 20 nt were selected as universal and type-specific primers, respectively.

The strategy for typing HCV is illustrated in Fig. 2. cDNA was prepared as outlined in Methods. The universal and four type-specific primers were expected to generate DNA fragments with a size specific for each of
the four HCV types. Type-specific PCR products were clearly recognized by their distinct sizes after they had been subjected to electrophoresis and stained with ethidium bromide (Fig. 3). When 39 isolates in the panel (clones no. 1 to 4, no. 7 to 26 and no. 30 to 44 in Fig. 1), for which samples were available, were typed by PCR, the results were in complete accord with the sequences corresponding to their type-specific primers.

A part of the putative C gene [nt 146 to 486 (341 bp), primer sequences excluded] was amplified in the 39 HCV isolates by PCR with universal primers (no. 23 and no. 122), and compared for the degree of sequence identity within and among the four types (Table 1). An identity of 97.4 to 95.1% was observed within each type, which was distinctly greater than that of 91.0 to 77.9% between any two different types. These results indicated that the sequences of 20 nt, adopted as type-specific primers, could be taken to represent the character of a longer sequence in the C gene.

**Various HCV types in blood donors and patients**

HCV samples from 159 blood donors, 80 patients with NANB liver disease and 11 haemophiliacs were typed by PCR with type-specific primers (Table 2). Some differences were noted in the distribution of HCV types between donors and patients. Although type II was most prevalent in both donors and patients with NANB liver disease in Japan, it was found more often (P < 0.01) in donors (131/159, 82%) than in patients (48/80, 60%). The prevalence of type III was lower in donors than in patients [16/159 (10%) compared to 18/80 (23%), P < 0.01]. HCV from 11 haemophiliacs was distributed in a somewhat different pattern, type I being found in five, and type II in four (mixed types included).

There were four cases in which HCV of two different types occurred together in the same sample, and therefore double infection with HCV strains of different types was suspected (Fig. 4). One of them was from a patient with chronic hepatitis (types I and II) and another from a patient with hepatocellular carcinoma (types II and III). The remaining two were from haemophiliacs (types I and II; I and III).

Double infection with HCV of different types was confirmed by sequence analyses within 341 bp in the C gene (nt 146 to 486). An identity of 90.1% was observed.

### Table 1. Degree of sequence identity in a part of the putative C gene within and among four different types of HCV

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of isolates</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>4</td>
<td>97.4 ± 0.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>20</td>
<td>91.0 ± 1.0</td>
<td>95.3 ± 1.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>III</td>
<td>10</td>
<td>79.5 ± 0.8</td>
<td>79.4 ± 1.1</td>
<td>95.1 ± 1.0</td>
<td>-</td>
</tr>
<tr>
<td>IV</td>
<td>5</td>
<td>79.1 ± 0.8</td>
<td>77.9 ± 0.7</td>
<td>85.5 ± 0.8</td>
<td>96.1 ± 0.6</td>
</tr>
</tbody>
</table>

* A total of 39 HCV isolates, classified into four types by PCR with type-specific primers (see Methods), were compared for sequence identity in 341 bp representing nt 146 to 486 in the C gene (primer sequences of PCR products at both ends were excluded from the comparison). Values represent mean ± s.d. percentage identity.

### Table 2. Prevalence of four HCV types in blood donors and patients with HCV infection

<table>
<thead>
<tr>
<th>Donor or patient type</th>
<th>Sample number</th>
<th>HCV type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood donors</td>
<td>159</td>
<td>I 6 II 131 III 16 IV 6 Mixed 0</td>
</tr>
<tr>
<td>Acute hepatitis</td>
<td>12</td>
<td>I 0 II 9 III 3 IV 0 Mixed 0</td>
</tr>
<tr>
<td>Chronic hepatitis</td>
<td>27</td>
<td>I 0 II 16 III 7 IV 3 Mixed 1</td>
</tr>
<tr>
<td>Liver cirrhosis</td>
<td>24</td>
<td>I 2 II 15 III 5 IV 2 Mixed 0</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>17</td>
<td>I 2 II 8 III 3 IV 1 Mixed 0</td>
</tr>
<tr>
<td>Haemophilia</td>
<td>11</td>
<td>I 3 II 3 III 2 IV 1 Mixed 0</td>
</tr>
</tbody>
</table>

* HCV in plasma or serum samples from blood donors, patients with various NANB liver diseases and haemophiliacs was typed by PCR with type-specific primers (see Methods).
Typing hepatitis C virus by PCR

Fig. 4. Two types of HCV occurring together in sera from patients with NANB liver disease or from haemophiliacs. Lane 1, a patient with chronic NANB hepatitis (types I and II); lane 2, a patient with NANB hepatocellular carcinoma (types II and III); lane 3, a haemophiliac (types I and II); lane 4, another haemophiliac (types I and III). Migration positions of PCR products of the four different HCV types are indicated on the right.

Fig. 5. Identical HCV types in mother and baby in two families and in donor and recipient after needle exposure. HCV isolates from baby and mother, in a family with perinatal transmission of HCV, are shown in lanes 1 and 2, and those from the other family in lanes 3 and 4, respectively. HCV from the donor of the accidental transmission is shown in lane 5, and that from the recipient in lane 6. Migration positions of PCR products of four different HCV types are indicated on the right.

Same HCV types in mother and baby with perinatal transmission, and in donor and recipient after needle exposure

HCV types were used to study perinatal transmission of HCV in two mother–baby pairs, and in a case of accidental HCV infection after needle exposure (Fig. 5). HCV isolates from both mother and baby in one family were typed as II, and those in the other as III. HCV isolates from both donor and recipient in needle exposure were typed as IV.

To establish further the route of infection, HCV isolates from the two mothers and babies, as well as those from the donor and recipient in the needle exposure, were sequenced in the nt 146 to 486 region of the C gene. In one mother–baby pair, the sequence of all 341 nt was identical; in the other, there was 99-4% agreement (339 of 341 nt). The sequence of all 341 nt was identical in HCV isolates from donor and recipient in the accidental transmission by needle.

Discussion

A method was developed for classifying HCV into four types, based on amplification by PCR with primers deduced from the putative C gene that were characteristic for each of them. PCR was carried out in two stages to increase the sensitivity of typing. The first PCR was performed on HCV cDNA with universal primers which generated fragments of 272 bp (nt 139 to 410 in the C gene). The second PCR was carried out with a universal, sense primer and a mixture of four type-specific, antisense primers. In order to increase the specificity of the reaction, annealing was performed at a higher temperature (60 °C) in the second PCR. By this method, distinct sizes of PCR products were expected for HCV types I, II, III and IV, i.e. 57, 144, 174 and 123 bp, respectively. This assumption was substantiated by correlating the results of typing with nucleotide sequences corresponding to type-specific primers in 39 HCV isolates. Further, HCV in all 256 serum samples tested was typed by this method; 252 showed one of the four types and the remaining four revealed two types. These findings would support the validity of this typing method. As precedents for such a method, dengue and hepatitis B viruses have been typed by PCR with typespecific primers (Deubel et al., 1990; Norder et al., 1990).

Of four reported isolates for which the entire sequence is known, HCV-1 (Choo et al., 1991) from the United States belongs to type I. Among the other three from Japan, HCV-J (Kato et al., 1990) and HCV-BK (Takamizawa et al., 1991) belong to type II, and HC-J6 (Okamoto et al., 1991) to type III. Enomoto et al. (1990) proposed four types of HCV based on sequence identity between two HCV isolates of distinct types from the patient with hepatitis, 80-3% between those from the patient with carcinoma, 88-8% between those from one haemophiliac and 79-4% between those from the other.
within 340 bp in the NS5 region of 10 isolates. By analysis of five reported isolates (Kato et al., 1990; Choo et al., 1991; Ogata et al., 1991; Okamoto et al., 1991; Takamizawa et al., 1991) and 12 unpublished ones, for which sequences of both the C gene and NS5 region are known, type I corresponds to PT, type II to K1, type III to K2a and type IV to K2b in their classification. Different HCV types infecting American and Japanese populations have been reported (Kato et al., 1990; Takeuchi et al., 1990; Choo et al., 1991). In Japan, HCV of type II was the most prevalent both in apparently healthy blood donors (82%o) and patients with NANB liver disease (60%o). In contrast, the prototype HCV (Choo et al., 1991), the H strain [estimated from the C gene and NS5 region sequences spanning 4923 bp (Ogata et al., 1991)] and three additional isolates from the United States (Ulrich et al., 1990) [estimated from 162 bp in the NS3 region based on the correlation of diversities between the C gene and NS3 region sequences (Okamoto et al., 1991)] were invariably of type I, as were 18 isolates from the United Kingdom [estimated from 216 bp in the NS3 region (Simmonds et al., 1990)].

There were some differences in the distribution of HCV types in blood donors and patients with acute or chronic NANB liver disease. Type II was found more often in blood donors than in patients (82% compared to 60%o, P < 0.01), whereas type III was found less often in donors than in patients (10% compared to 23%o, P < 0.01). The prevalence of type II HCV was the lowest (47%) in 17 patients with hepatocellular carcinoma. It is not certain, however, whether the observed differences reflect a stronger propensity of type III HCV to induce clinical disease than type II. The association of HCV types with clinical disease would have to be evaluated in an extended series of patients and symptom-free carriers, and in various geographical areas where different HCV types are prevalent.

HCV of type I was frequently observed in haemophiliacs in Japan (5/11) who had received coagulation factor concentrates, imported from the United States in the distant past, which had not been heat-inactivated and were possibly contaminated with HCV (Garson et al., 1990). Based on a sequence identity within 275 bp in the NS5 region, Hijikata et al. (1990) also identified HCV of the US strain (type I in our classification) in three of four haemophiliacs in Japan.

Double infection with HCV of different types was detected in four of 256 HCV-positive samples. They were from two patients with chronic NANB liver disease and two haemophiliacs. Co-occurrence of HCV of two distinct types was confirmed by 10 to 21% differences in a sequence of the C gene [nt 146 to 486 (341 bp)]. There is a possibility that the prevalence of double infection would have been underestimated by the present typing method that employs a mixture of four antisense primers in the second PCR. It would be necessary to perform the second PCR with each of four type-specific primers in serum samples when double infection is suspected.

How HCV is transmitted and maintained within the community has not yet been established. Alter et al. (1990) reported that percutaneous transmission, such as transfusion or illicit drug use, would account for 50% of cases at most. Perinatal and intra-familial transmissions (Reesink et al., 1990; Tajima et al., 1991), as well as spread through sexual contacts (Alter et al., 1989b), would be expected to fill some of this gap.

Perinatal HCV transmission was detected by demonstrating the same HCV type in mother and baby in each of two families. The HCV isolates possessed identical sequences within 341 bp in the C gene, or showed a 99-4% concordance, providing further proof for the route of infection; even HCV isolates of the same type were identical in 95-1 to 97-4% within this sequence (Table 1).

In addition, the source of transmission was confirmed in a case of accidental needle exposure, in which HCV specimens from both donor and recipient were of type IV. HCV typing would be particularly useful in determining the route of infection when more than one source is suspected.

References


(Received 9 September 1991; Accepted 7 November 1991)