cDNA clones of Japanese hepatitis C virus genomes derived from a single patient show sequence heterogeneity

M. Oshima,1 M. Tsuchiya,1 M. Yagasaki,1 T. Orita,1 M. Hasegawa,1 K. Tomonoh,1 T. Kojima,1 Y. Hirata,1 O. Yamamoto,1 Y. Sho,2 E. Maeda2 and T. Arima2

1Fuji-Gotemba Research Laboratories, Chugai Pharmaceutical Co. Ltd, 1-135 Komakado, Gotemba-shi, Shizuoka 412 and 2The Second Department of Internal Medicine, Faculty of Medicine, Kagoshima University, 1208-1 Usuki-cho, Kagoshima 890, Japan

Twelve cDNA clones of Japanese hepatitis C virus (HCV) have been isolated from liver tissue of a single non-A, non-B hepatitis patient. These clones represented the non-structural domains of HCV. The degree of substitution in the nucleotide sequences and deduced amino acid sequences between these clones was 9.5 and 7.7%, respectively. This high level of substitution suggested that repeated infections of different HCVs may have occurred in the patient.

Most cases of post-transfusion hepatitis have been thought to be caused by non-A, non-B hepatitis (NANBH) virus (Alter et al., 1975). A cDNA of the NANBH virus genome has recently been cloned in the U.S.A. (Choo et al., 1989) from the plasma of chimpanzees chronically infected with NANBH agents. The nucleotide sequence of the clone has also been determined (Houghton et al., 1989). This virus has been named hepatitis C virus (HCV) (Choo et al., 1989). HCV, a positive-stranded RNA virus, is thought to be related to the Flaviviridae (Choo et al., 1989).

In Japan, cDNA clones which cover the whole HCV genome have also been cloned from plasma of NANBH patients by using the polymerase chain reaction (PCR), or from a cDNA library using the method of Benton & Davis (1977) (Kato et al., 1990a). The nucleotide sequences of HCV cDNA clones isolated in Japan (HCV-J) are 20 to 30% different from the sequence originally cloned in the U.S.A. (HCV-US), indicating the existence of a Japanese-type HCV. Furthermore, even in the cDNA clones derived from Japanese-type HCV, a 2.5 to 11% diversity was observed in the nucleotide sequences (Kato et al., 1990a, b), suggesting the existence of variants or subtypes of HCV.

Heterogeneity of the HCV genome in a single patient has not yet been reported. In the present paper, we describe the isolation and characterization of 12 HCV clones (HCV-22) which were isolated from a λgt11 cDNA library constructed with liver RNA from a single NANBH patient.

Total RNA was extracted according to the acid guanidinium thiocyanate–phenol–chloroform method (Chomczynski & Sacchi, 1987) from 1 g of liver tissue obtained from a single Japanese NANBH patient (a 70-year-old surgeon) during surgical resection of a hepatoma. Double-stranded cDNAs were synthesized according to Gubler & Hoffmann (1983) using 15 μg of RNA preparation as a template with random hexanucleotide primers. These cDNA fragments were cloned into λgt11 (Young & Davis, 1983) after addition of an EcoRI linker and digestion with EcoRI. The recombinant λgt11 was packaged in vitro. The λgt11 cDNA library, containing 1.5 × 10⁸ p.f.u., was immunoscreened according to the method of Young & Davis (1983).

The initial immunoscreening of the library with 20-fold diluted mixed sera from five acute convalescent and five chronic NANBH patients gave 29 positive clones. These clones were further examined with a serum panel obtained by Dr H. J. Alter (NIH, Bethesda, Md., U.S.A.) in order to select NANBH-specific clones. This panel contained 26 independent serum samples collected from 13 persons, including three NANBH patients and two implicated donors with NANBH. Finally, 15 clones were obtained as NANBH-specific clones by selection with this panel. The other 14 clones were thought to be non-specific because they reacted with some sera from normal donors and/or other liver disease patients. Table 1 shows the immunoreactivity of the 15 NANBH-specific clones with the serum panel. The cDNA fragments of these clones were recovered from recombinant λgt11 DNAs by double digestion with KpnI and SacI, and subcloned into the KpnI and SacI sites of pUC18 or pUC19. The cutting sites of KpnI and SacI in the λgt11 DNA were positioned about 1 kb from...
Table 1. Serological test with serum panel

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2232</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2230</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2206</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2258</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2248</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2216</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2246</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2211</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2256</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2207</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2244</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2239</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2270</td>
<td>++</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>++</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2264</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

* Samples are categorized as follows. A/N, E/R, chronic transfusion-associated NANBH; C/P, I/V, J/W, implicated donor with chronic NANBH; F/S, acute TA NANBH; D/Q, L/Y, liver disease control; B/O, G/T, H/U, K/X, M/Z, normal donor. Clones were tested in duplicate.

The EcoRI site in the 5' and 3' directions, respectively. Therefore, the subcloned cDNA inserts included about 1 kb of λgt11 DNA at both ends. The nucleotide sequences of both strands of cDNA of these isolated clones were determined by the dideoxynucleotide chain termination method (Messing, 1983) using λgt11 primers.

Three clones, 2239, 2264 and 2270, which showed lower immunoreactivity than the other clones (Table 1), were about 200 bp in length. These had no homology with any nucleotide sequences in GenBank and no homology with each other. The other 12 clones were derived from the HCV genome and were located between nucleotide positions 4126 and 6026 (using the numbering of Houghton et al., 1989), corresponding to the non-structural protein 4 (NS4) and 5 (NS5) domains. These 12 HCV clones were collectively termed HCV-22. All clones of HCV-22 had been cloned into the EcoRI site of λgt11 using the endogenous EcoRI site of HCV at nucleotide position 5010. HCV-22 could thus be divided into two groups with nucleotide position 5010, the endogenous EcoRI site, as a boundary point. Fig. 1 shows the location of HCV-22 in the HCV genome. The nucleotide sequences of HCV-22 showed 74 to 77% identity with the sequence of HCV-US (Houghton et al., 1989) and 86 to 88% identity with the sequence of HCV-J (Kato et al., 1990a).

The nucleotide sequences at positions 4416 to 5015 (600 bp) and 5010 to 5609 (600 bp) of HCV-22 were compared as shown in Fig. 2(a) and (b). Their corresponding deduced amino acid sequences were compared as shown in Fig. 3(a) and (b). Clone 2216 has not been included in the comparison because of base insertions at two positions, which resulted in a frameshift. Although base substitutions in the nucleotide sequences among HCV-22 occurred at 150 different positions, the substitutions in the amino acid sequences only occurred at 29 positions. This indicates that 81% of the nucleotide positions in which base changes were observed were silent mutations. This finding supports the previously reported claim that numerous silent mutations occur in the HCV genome (Kubo et al., 1989; Enomoto et al., 1990; Takeuchi et al., 1990). The sequences were almost identical between clones 2207 and 2244, 2206 and 2258, 2248 and 2220, and 2232 and 2230. The sequence diversity of the nucleotide and the amino

---

Fig. 1. The regions of HCV-22 on the HCV genome compared with HCV-US (Houghton et al., 1989). The sequence numbering is the same as that of Houghton et al. (1989). These clones are located over NS4 and NS5 which cover 1900 bp of the HCV genome, and are divided by the endogenous EcoRI site (at nucleotide position 5010).
acid sequences was less than 1.3 and 0.6%, respectively, in each pair. However, the degree of substitution of the amino acid sequences between all clones of HCV-22 was 9.5 and 7.7%, respectively. Although the base changes were observed throughout the nucleotide sequences of HCV-22, the amino acid sequences were
completely conserved over 30 residues in some regions (see Fig. 3). Conversely, frequent amino acid substitutions (6/11) were observed in positions 158 to 168 (using the numbering in Fig. 3).

Nucleotide sequences for the whole HCV genome have been independently obtained in Japan (HCV-J by Kato et al., 1990a) and in the U.S.A. (HCV-US by Houghton et al., 1990). Other groups (Kubo et al., 1989; Maeno et al., 1990) reported the nucleotide sequences of HCV clones encoding the two non-structural regions derived from Japanese patients which showed 7 to 10% diversity from HCV-J. On the other hand, Enomoto et al., (1990) reported that there are two major types of HCV in Japan which show only 67% similarity in the nucleotide sequences with each other. The nucleotide sequences of HCV-22 show 11 to 13% diversity from that of HCV-J. It is suggested that all clones of HCV-22 and HCV-J are variants of the same type of HCV. All these cDNA clones isolated in Japan show 20 to 30% diversity from HCV-US, demonstrating that Japanese-type HCV is different from HCV-US.

In this study, heterogeneity of the HCV genome was found even though the HCV originated from only one patient. Kato et al. (1989) and Kubo et al. (1989) have isolated HCV cDNA clones by PCR from a patient and a blood donor, and have shown that the nucleotide substitutions occurred in about 1% of their isolates. They have suggested the possibility of mutation of the HCV genome in a single patient. However, they isolated their cDNA clones by the PCR method. We obtained the HCV-22 clones by immunoscreening, in which errors due to Taq polymerase during the PCR and the selection of cDNA with sequences of the PCR primers used are negligible. We found up to 9.5 and 7.7% substitution in nucleotide and amino acid sequences, respectively, among the HCV-22 clones. Such a high degree of variation of HCV in a single patient has not been previously reported. Nucleotide substitution of less than 10% was observed between four pairs of HCV-22 (see Fig. 2) and may have arisen in the patient during the long incubation periods after infection. However, it seems less likely that all variants of HCV-22 were derived from mutations of a single HCV in the patient, as flaviviruses show a higher gene stability than other RNA viruses (Rice et al., 1986). For example, some strains of dengue viruses isolated 20 years apart have been found to have a low frequency of mutation (Trent et al., 1983). Additionally, Kato et al. (1990b) have reported that the nucleotide sequences of their HCV clones derived from different patients showed 3% diversity. Therefore, the HCV-22 variants which show a high degree of nucleotide substitution are suggested to have been derived from different infectious sources. In this case, the patient is a 70-year-old surgeon who could have been exposed to the risk of HCV infection during numerous surgical operations. It is possible that NANBH patients who have received repeated blood transfusions may also have been infected by HCV repeatedly.

Our data should provide valuable information towards the development of an HCV vaccine or other therapeutic agents for HCV, although many problems remain to be resolved.

Fig. 3. Deduced amino acid sequence alignments of HCV-22. Parts (a) and (b) correspond to Fig. 2(a) and (b), respectively.

References


(Received 3 April 1991; Accepted 31 July 1991)