Typing of hepatitis C virus genomes by restriction fragment length polymorphism

Toshifumi Nakao,¹ Nobuyuki Enomoto,² Nobuo Takada,² Akira Takada² and Takayasu Date^{1*}

¹Division of Cancer Research, Medical Research Institute and ²Division of Gastroenterology, Department of Internal Medicine, Kanazawa Medical University, Uchinada, Ishikawa 920-02, Japan

Recently, we reported that hepatitis C virus (HCV) can be classified genetically into two types, HCV-K1 and HCV-K2, which show 67% and 71% identity at the nucleotide and amino acid sequence levels in a 340 bp region which encodes the NS5 gene Gly-Asp-Asp motif. To develop a rapid method to classify the genomes of HCV isolates, we identified restriction fragment length polymorphisms (RFLPs) in reverse transcriptase–polymerase chain reaction products encoding a portion of the NS5 gene. *AluI* and *AccII* enabled HCV to be classified into the K1 and K2 types, and *Sau96I* enabled classification into the K1 type, and the K2a and K2b subtypes. These RFLPs also generally

allow Japanese isolates to be distinguished from the prototype (PT, an isolate from the U.S.A.), which is a K1 type. Sequence analysis of the 5'-untranslated regions of Japanese isolates revealed near identity between the K1 type and PT, and 93 to 94% identity between the K1 and K2 types, indicating that there are type K1- and K2-specific RFLPs in this region. Our results suggest that the nucleotide sequences of the K1 and K2 types are different throughout the HCV genome. The incidence of HCV types K1, K2a and K2b, and PT in 50 samples was 74%, 16%, 8% and 2%, respectively.

Introduction

Hepatitis C virus (HCV) is the major cause of post-transfusion non-A, non-B hepatitis (Kuo *et al.*, 1989) and is considered to be related to the *Flaviviridae*. HCV cDNA was initially cloned from the plasma of a chimpanzee chronically infected with non-A, non-B hepatitis (Houghton *et al.*, 1988; Choo *et al.*, 1989). The first genome type, isolated in the U.S.A. is designated here as the prototype (PT). Subsequently, two groups have cloned and sequenced HCV genomes from Japanese non-A, non-B hepatitis patients (Kato *et al.*, 1990; Takamizawa *et al.*, 1991). The HCV genome is about 9.4 kb long, single- and positive-stranded, and does not have a poly(A) tract at the 3' terminus. The genome encodes two putative structural proteins which are translated from sequences in the vicinity of the 5' terminus of the genome: nucleocapsid protein, containing a number of basic amino acids, and envelope protein, containing nine potential glycosylation sites. Non-structural proteins NS1 to NS5 are translated from the coding region immediately following the envelope gene (Okamoto *et al.*, 1990a; Kato *et al.*, 1990; Takeuchi *et al.*, 1990a; Takamizawa *et al.*, 1991).

Molecular cloning and application of the reverse transcription–polymerase chain reaction (RT–PCR) to

the HCV genome revealed that the nucleotide sequences of Japanese strains are slightly different from that of PT (Kato *et al.*, 1989; Kubo *et al.*, 1989; Kaneko *et al.*, 1990; Maeno *et al.*, 1990; Takeuchi *et al.*, 1990b,c), and that there are HCV variants in Japan (Enomoto *et al.*, 1990). The nucleotide sequences of the structural and non-structural genes of the American and Japanese strains differ by as much as 34%; in contrast, a short 5' non-coding region, which may function as a replication origin, is conserved extensively among four Japanese isolates and PT (Okamoto *et al.*, 1990a). However, it is not known whether the nucleotide sequence of the non-coding region is highly conserved among all HCV strains.

Traditionally, viruses have been classified into serotypes on the basis of their antigenic characteristics, but with the recent development of molecular biology it is possible to classify viruses genotypically. Recently, we reported the nucleotide sequence of a portion of the NS5 gene from HCV RNA isolated from 10 Japanese individuals, and demonstrated that HCV can be classified into at least two types, HCV-K1 and HCV-K2, which show 67% and 71% identity at the nucleotide and amino acid sequence level, respectively (Enomoto *et al.*, 1990). Slot blot hybridization indicated that PT is related to HCV-K1, the major type in Japan. Furthermore,

HCV-K2 is divided into two subtypes, HCV-K2a and HCV-K2b, which show 80% nucleotide sequence identity.

However, it is not known whether all Japanese isolates belong to the types mentioned above whether specific virus genotypes are responsible for certain forms of the disease, or for particular widespread epidemics. Therefore, important information concerning the detection of virus in transfusion patients' blood and strategies for disease therapy should come from an understanding of the genetic variability and world-wide distribution of HCV.

In this study, we have compared the typing of HCV by slot blot hybridization analysis and restriction fragment length polymorphism (RFLP) using three restriction endonucleases to analyse RT-PCR products. Our results demonstrated that HCV cDNAs show type-dependent RFLPs which can be used to classify isolates according to type or subtype.

Methods

Viruses. Fresh plasma was obtained from 50 patients with chronic non-A, non-B hepatitis; 12 had chronic hepatitis, 15 had acute post-transfusion hepatitis, 20 had liver cirrhosis and three had hepatocellular carcinoma. The frequency of detection of a 401 bp fragment was dependent on the disease: chronic hepatitis, 93.5%; acute hepatitis, 28.6%; liver cirrhosis, 81%; hepatocellular carcinoma, 33.3% (Enomoto *et al.*, 1991).

Viral RNA preparation. Viral RNA was extracted from 1 ml of plasma using acid-guanidinium thiocyanate-phenol-chloroform (Chomczynski & Sacchi, 1987) and the final product was dissolved in 20 µl water. To amplify the 401 bp fragment, an aliquot of RNA (2 µl) was mixed with 10 µl buffer containing 10 pmol AS-401 primer (Table 1), 50 mM-Tris-HCl pH 8.3, 75 mM-KCl, 3 mM-MgCl₂ and 10 mM-DTT, and heat-treated for 5 min at 70 °C. After quick chilling, 20 units (U) of RNase inhibitor and 100 U of Moloney murine leukaemia virus reverse transcriptase (Promega) were added, and the mixture was incubated for 45 min at 37 °C and 5 min at 100 °C. To amplify the 225 and 242 bp fragments, 50 pmol random hexamer primer (Takara Biochemicals) was used as a primer for reverse transcriptase. The random primer produced cDNA more efficiently than the defined 3' primers used for amplification.

Amplification by PCR. PCR was carried out in a DNA thermal cycler

(Perkin-Elmer Cetus) for 50 cycles using the method of Saiki *et al.* (1988). Each reaction cycle included denaturation at 94 °C for 1 min, and primer annealing at 40 °C and primer extension at 72 °C for 1 min each. The sequences of oligonucleotide primers, synthesized on an Applied Biosystems 381A machine, are shown in Table 1. Aliquots of products (10 µl) were analysed by 4% PAGE and 1 µl was used for slot blot analysis. For the RFLP analysis, PCR products were separated on a 4% polyacrylamide gel, and regions corresponding to the 401 bp fragment were excised and immersed in 100 µl 10 mM-Tris-HCl pH 8.3 and 0.1 mM-EDTA for 6 h at 50 °C. Extracted DNA (1 µl) was subjected to a second PCR, and products were precipitated with ethanol and used for RFLP analysis.

For slot blot hybridization analysis a portion (1/100) of each PCR reaction was mixed with 200 µl of 0.5 M-NaOH, 1.5 M-NaCl solution, divided into two 100 µl aliquots and subjected to slot blot hybridization on a nylon filter (Pall; Biotryne). Each pair of duplicate filters was incubated with 20 ng of digoxigenin-labelled HCV-K1 or HCV-K2a 401 bp fragment probe, as described previously (Enomoto *et al.*, 1990). The hybridized, digoxigenin-labelled probe was visualized with an immunostaining kit (Boehringer Mannheim).

RFLP experiments. RT-PCR products were precipitated using ethanol and then solubilized with 10 µl TE buffer. Aliquots (3 µl) were taken and digested with 2 U of restriction endonuclease in 20 µl under optimal conditions for 3 h. Digested samples (5 µl) were mixed with 3.5 µl of sample buffer, separated on a 6% polyacrylamide gel and stained with 1 µg/ml ethidium bromide.

Nucleotide sequencing. Nucleotide sequences of both plus and minus strands were determined directly by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) using PCR products generated by asymmetric amplification with unequal molar amounts of primers (Gyllenstein & Erlich, 1988).

Results

Amplification and typing of the 401 bp fragment

Amplification from non-A, non-B hepatitis patient plasma of the HCV cDNA 401 bp fragment, which encodes a portion of the NS5 gene, was carried out by RT-PCR using S and AS primers. HCV genomes were detected with high frequency from chronic hepatitis patients (Enomoto *et al.*, 1991). Amplified products from 50 patients were classified by slot blot hybridization using HCV-K1 or HCV-K2a digoxigenin-labelled probes; 38 hybridized with the HCV-K1 probe, whereas 12 samples hybridized with the HCV-K2a probes. All

Table 1. PCR primers

Name	Sequence (5' to 3')	Location of sequence amplified	Fragment size (bp)	Reference
S	TGGGGATCCCGTATGATACCCGCTGCTTTGA	NS5 gene	401	Enomoto <i>et al.</i> (1990)
AS	GGCGGAATTCCTGGTCATAGCCTCCGTGAA			
SS'	ACTCCACCATAGATCACTCC	5'-Untranslated region	242	Kanai <i>et al.</i> (1990)
AS5'	AACACTACTCGGCTAGCAGT			
S'	TGCGGTTATTGCCGTTGTCGCGCCAGCGG*	NS5 gene	225	This work
AS'	GGCAGAATACCTAGTCATGGCCTCTGTGAA*			

* The binding sites of primers S' and AS' are within the region amplified by the S and AS primers.

Fig. 1. Restriction endonuclease digestion patterns of the 401 bp RT-PCR cDNA fragment of HCV which encodes a portion of the NS5 gene. Amplified cDNA fragments were digested with three different restriction endonucleases and subjected to 6% PAGE. Gels were stained with ethidium bromide and photographed on a U.V. light transilluminator. (a) *AluI* digests; (b) *AccII* digests; (c) *Sau96I* digests. Samples are indicated at the top; only 12 HCV-K1 samples are shown. Lane M contains DNA size markers, a *HincII* digest of ϕ X174 DNA.

samples were hybridization-positive, but none was positive for both probes, suggesting that there is no other type of HCV and no intermediate type in these samples.

Of 50 samples, 10 (five K1, one PT, two K2a and two K2b) had been sequenced previously and their nucleotide sequences reported (Enomoto *et al.*, 1990). On the basis of these sequence data, restriction enzyme sites which may have been useful for classifying each HCV genotype were identified by computer analysis; *AluI*, *AccII* and *Sau96I* were selected and used for RFLP analysis. Samples were divided into K1, K2a and K2b type groups on the basis of hybridization and sequence analysis (see below).

Classification of K2 type

Since slot blot hybridization was unable to classify PCR products into K2 subtypes, the 401 bp fragment of all K2 type samples was sequenced. Based on this analysis, all the samples could be subtyped; eight were K2a and four

were K2b. Variation between the sequences of members of subtypes K2a and K2b was less than 7% and 8%, respectively, and both subtypes had 66 to 69% identity with members of the K1 group.

RFLP analysis

All 50 of the 401 bp HCV cDNA fragments were digested with *AluI* and the resultant fragments were separated by PAGE. *AluI* digestion of 37 K1 type samples produced characteristic 220 and 123 bp fragments, accompanied by either a 58 or 50 bp fragment; two K1 type samples showed a different pattern (Fig. 1a). In contrast to type K1, *AluI* digestion of type K2 samples generated four different patterns with no common fragments. Digestion of PT with *AluI* resulted in 220 and 181 bp fragments, and its RFLP pattern was different from those of types K1 and K2. Thus, the *AluI*-generated RFLPs are very useful for the classification of the 401 bp fragment into types K1 and K2, and probably

Table 2. *AluI*-, *AccII*-, and *Sau96I*-generated RFLPs of the 401 bp cDNA fragment

RFLP*	HCV type or subtype	Incidence†	Fragment (bp)‡
<i>AluI</i> digestion			
Aluk1	K1	35/37	220, 123, 58 or 50
Aluk1-1	K1	1/37	175, 131, 50, 45
Aluk1-2	K1	1/37	128, 123, 92, 50
Alupt	PT	1/1	220, 181
Aluk2-1	K2	4/12	261, 140
Aluk2-2	K2	3/12	401
Aluk2-3	K2	3/12	270, 99
Aluk2-4	K2	2/12	369
<i>AccII</i> digestion			
Acck1	K1	25/37	196, 140, 65 or 47
Acck1-1	K1	3/37	196, 158, 47
Acck1-2	K1	2/37	140, 138, 58, 47
Acck1-3	K1	2/37	336, 65
Acck1-4	K1	1/37	196, 77, 63, 47
Accpt	PT	—	248, 106, 47§
Accpt-1	PT	1/1	158, 106, 90, 47
Acck2	K2	4/37	205, 196
	K2	10/12	
Acck2-1	K2	1/12	205, 120, 76
Acck2-2	K2	1/12	196, 128, 77
<i>Sau96I</i> digestion			
Sauk1	K1	27/37	253, 93, 55
Sauk1-1	K1	5/37	253, 148,
	K2b	1/4	
Sauk1-2	K1	4/37	193, 93, 60
Sauk1-3	K1	1/37	213, 93, 55
Saupt	PT	—	147, 105, 90, 59
Saupt-1	PT	1/1	104, 93, 90
Sauk2a-1	K2a	3/8	233, 101
Sauk2a-2	K2a	3/8	233, 79, 68
Sauk2a-3	K2a	1/8	233, 147
Sauk2a-4	K2a	1/8	173, 102, 60
Sauk2b	K2b	3/4	148, 136, 117

* RFLPs were numbered in the order of frequency of occurrence. Those without a hyphen and suffixed number are the major RFLPs (more than 60%) for each type or subtype.

† The incidence is the number of occurrences of an RFLP type/total number of samples of each type or subtype.

‡ Fragments smaller than 40 bp were not evaluated.

§ Predicted fragments deduced from the sequence of PT are presented because its RFLPs are different from those of PT-1 (Accpt-1).

PT, which contains a unique 181 bp fragment. The major RFLP type of type K1 is designated Aluk1; the minor RFLP patterns were numbered in order of frequency. *AluI*-generated RFLP patterns and their frequencies are summarized in Table 2.

The *AccII* RFLPs of the same samples were analysed (Fig. 1b and Table 2). In contrast to the *AluI*-generated RFLPs, type K1 samples showed six different RFLPs; the majority (25 of 37) contained 196 and 140 bp fragments with either a 65 or 47 bp fragment (this RFLP was designated Acck1); eight other type K1 samples also contained a 196 bp fragment. One RFLP, designated Acck1-4, was found only once. On the other hand, most type K2 samples (10 of 12) showed a common RFLP

Fig. 2. Nucleotide sequences of amplified 5'-untranslated regions of types K1, K2a and K2b-1, and PT-1. Fragments were amplified by RT-PCR using the primers described. The sequences of HCV J1 and J4 are from Okamoto *et al.*, (1990a). *HaeIII* (H) and *Sau3A* (S) sites are shown above the nucleotide sequences; parentheses indicate predicted endonuclease sites found in one subtype K2a fragment.

consisting of 205 and 196 bp fragments. Interestingly, four type K1 samples showed the same digestion pattern, precluding classification by *AccII*-generated RFLP. The RFLPs of PT-1 and PT were different, and neither was found in those of Japanese strains.

The *Sau96I* digestion of the 401 bp fragment of most type K1 samples (32 of 37) produced a characteristic 253 bp fragment, whereas from seven of 12 of the type K2 samples a 233 bp fragment was largest (Fig. 1c and Table 2). Sequence analysis indicated that all PCR products containing a 233 bp *Sau96I* fragment were of subtype K2a, and three of four subtype K2b samples produced identical patterns containing three unique fragments. This indicates that *Sau96I* may be useful for classifying HCV into the K2a and K2b subtypes. One of the subtype K2b samples showed a *Sauk1*-1 type RFLP, but because its nucleotide sequence was 96% identical to that of K2b, it is clearly a member of the K2b subtype.

Nucleotide sequence and RFLP in the 5'-terminal region

Okamoto *et al.* (1990a) have reported that the 5'-terminal (untranslated) region of the HCV genome is highly conserved (almost 99%) between two distantly related isolates, which facilitates the diagnosis of non-A, non-B hepatitis by PCR (Okamoto *et al.*, 1990b; Kanai *et al.*, 1990). By using the primers of Okamoto *et al.* we attempted to amplify a 242 bp fragment from our samples to investigate whether this sequence is conserved. As shown in Table 3, these primers enabled the amplification of a 242 bp fragment from the HCV-K1 genome. Nucleotide sequence analysis showed 99% identity between K1, PT-1 and other reported HCV fragments (J1 and J4), and 93 to 94% identity between types K1 and K2 (Fig. 2). These results suggest that type K1 and PT are closely related, whereas types K1 and K2

Table 3. RFLPs of the 401 bp and the 242 bp HCV cDNA fragments

HCV type	Nucleotide sequence	Hybridization	RFLP in NS5			RFLP in 5'-UT*	
			<i>AluI</i>	<i>AccII</i>	<i>Sau96I</i>	<i>HaeIII</i>	<i>Sau3A</i>
K1	K1	K1	k1	k1	k1	k1	k1
K1-1	K1	K1	k1	k1-4	k1	k1	k1
K1-2	K1	K1	k1	k1	k1		
K1-3	K1	K1	k1	k1-3	k1-1		
K1-4	K1	K1	k1	k1-1	k1		
K1-5		K1	k1	k1	k1	k1	k1
K1-6		K1	k1	k2	k1		
K1-7		K1	k1	k1-2	k1	k1	k1
K1-8		K1	k1	k1-2	k1	k1	k1
K1-9		K1	k1	k1-1	k1-2		
K1-10		K1	k1	k1	k1-3†	k1	k1
K1-11		K1	k1-1†	k1	k1		
K1-12		K1	k1-2†	k1	k1		
K1-13		K1	k1	k1	k1		
K1-14		K1	k1	k1	k1		
K1-15		K1	k1	k2	k1		
K1-16		K1	k1	k2	k1		
K1-17		K1	k1	k1-3	k1		
K1-18		K1	k1	k1	k1		
K1-19		K1	k1	k1	k1	k1	k1
K1-20		K1	k1	k1	k1		
K1-21		K1	k1	k1	k1		
K1-22		K1	k1	k1	k1	k1	k1
K1-23		K1	k1	k1	k1-1		
K1-24		K1	k1	k1	k1-1		
K1-25		K1	k1	k1	k1-2		
K1-26		K1	k1	k1	k1-1		
K1-27		K1	k1	k1	k1-1		
K1-28		K1	k1	k1	k1		
K1-29		K1	k1	k1	k1		
K1-30		K1	k1	k2	k1		
K1-31		K1	k1	k1	k1		
K1-32		K1	k1	k1	k1		
K1-33		K1	k1	k1	k1		
K1-34		K1	k1	k1	k1		
K1-35		K1	k1	k1	k1-2	k1	k1
K1-36		K1	k1	k1-1	k1-2		
PT-1	PT	K1	pt	pt-1	pt-1	k1	k2pt
K2a	K2a	K2	k2-4	k2	k2a-3	k2	k2pt
K2a-1	K2a	K2	k2-1	k2	k2a-1	K2‡	k2pt
K2a-2	K2a	K2	k2-1	k2	k2a-2	-§	-
K2a-3	K2a	K2	k2-2	k2	k2a-2	-	-
K2a-4	K2a	K2	k2-1	k2	k2a-2	-	-
K2a-5	K2a	K2	k2-2	k2	k2a-1	-	-
K2a-6	K2a	K2	k2-1	k2a-2	k2a-1	-	-
K2a-7	K2a	K2	k2-2	k2	k2a-4†	-	-
K2b	K2b	K2	k2-3	k2	k2b		
K2b-1	K2b	K2	k2-4	k2	k2b	k2	k2pt
K2b-2	K2b	K2	k2-3	k2	k2b	k2‡	k1
K2b-3	K2b	K2	k2-3	k2b-1†	k1-2	-	-

* 5'-UT, 5'-Untranslated region.

† Rare RFLP observed in only one case.

‡ Additional *HaeIII* sites were observed.

§ -, No 242 bp fragment was observed after RT-PCR.

are more distantly related, as we proposed previously (Enomoto *et al.*, 1990).

These amplified fragments were digested with *HaeIII* or *Sau3A* and subjected to PAGE to identify RFLPs. Fig. 3 shows that no type K1-derived fragment has an

HaeIII site, but that subtypes K2a and K2b have one conserved *HaeIII* site, producing 147 and 94 bp fragments and two of four contain additional *HaeIII* sites, e.g. one subtype K2a 242 bp fragment contains two additional *HaeIII* sites. On the other hand, *Sau3A*

Fig. 3. *Hae*III (a) and *Sau*3A (b) digestion of the 242 bp fragment from the 5'-untranslated region cDNA fragment of HCV. Amplified PCR fragments were digested with either *Hae*III or *Sau*3A and subjected to 6% PAGE. Lanes 1 to 8, samples K1, K1-1, K1-2, PT-1, K2a, K2a-1, K2b-1 and K2b-2.

digestion of the type K1 fragments and one of the subtype K2b-derived fragments generated two fragments of 178 and 63 bp, whereas digestion of PT-1 and other type K2-derived fragments generated a 228 bp fragment (the S5' primer contains a *Sau*3A site). These preliminary results indicated that *Hae*III digestion of the 242 bp 5'-terminal region can probably be utilized to classify HCV into HCV types K1 (including PT) and K2.

Discussion

We have shown previously that the 401 bp cDNA fragment containing a portion of the NS5 gene of HCV is amplified with high frequency from the plasma of non-A, non-B hepatitis patients by RT-PCR using the S and AS primers, and we proposed that there are two major types of HCV in Japan (Enomoto *et al.*, 1990). In this study we collected more non-A, non-B hepatitis samples and determined their HCV type by slot blot hybridization analysis, digestion with three restriction endonucleases and nucleotide sequencing. Our results provide further evidence that only two major types of HCV exist in Japan, HCV-K1 and HCV-K2; HCV-K2 can be

further subdivided into two subtypes, HCV-K2a and HCV-K2b. In this analysis, neither an intermediate nor a new type of HCV were found. The American type designated PT, which is a member of the K1 type family by slot blot hybridization and nucleotide sequence analyses, was found only in one non-A, non-B hepatitis patient, infected with virus from factor VIII preparation imported from the U.S.A. The RFLPs of the 401 bp fragment generated using three restriction enzymes are summarized in Table 3. *Alu*I-generated RFLPs seem to be most useful in the classification of HCV into types K1 or K2, whereas *Sau*3A-generated RFLPs facilitate the division of type K2 into its subtypes. Digestion with one restriction enzyme may generate a unique RFLP, but analysis with the other two restriction enzymes produces major or minor RFLPs and therefore, with relatively few exceptions, Japanese variants of HCV can be classified by RFLP analysis with three restriction endonucleases.

Recently, two groups have reported the complete nucleotide sequences of HCV genomes isolated from Japanese patients (Kato *et al.*, 1990; Takamizawa *et al.*, 1991). Both HCV genomes belong to the K1 type, based on the sequence of a portion of the NS5 gene; one (that of Kato *et al.*) shows Aluk1, Acck1-1, and Sauk1-3 RFLPs, and the other (that of Takamizawa *et al.*) shows Aluk1, Acck1 and Sauk1 RFLPs. Furthermore, both 5'-terminal regions contain *Hae*III-k1 and *Sau*3A-k1 RFLPs, indicating the usefulness of RFLP analysis.

Since the predicted amino acid sequence encoded by the 401 bp fragment contains the Gly-Asp-Asp motif conserved in most RNA viruses (Argos, 1988; Pletnev *et al.*, 1990), it is likely that amino acid sequences in other portions of the genome of types K1 and K2 will show less identity. It is not known whether these types contain characteristic amino acid and nucleotide sequences in portions other than the NS5 gene region tested or whether the failure to amplify other regions is due to differences in the HCV type. Amplification and cloning of type K2 cDNA containing other protein coding regions has been attempted, but no cDNA was obtained. However, a 242 bp fragment from the 5'-untranslated region has been amplified from plasma infected with different HCV types. The nucleotide sequences of subtypes K2a and K2b and PT-1 are 94%, 93%, and 99% identical to that of type K1, respectively, and the 242 bp fragment also showed type K1- and K2-dependent RFLPs. Interestingly, all 12 type K1 samples direct amplification of the 242 bp fragment and produce the same *Hae*III- and *Sau*3A-generated RFLPs, but only four of 11 type K2 samples produce the 242 bp fragment. The failure to amplify the 242 bp fragment from several type K2-containing plasma samples is probably due to nucleotide variation in the 3' primer-binding region. Although information on the 5' region is limited, these

results support our previous proposal that type K1 is genetically related to PT whereas type K2 is more distantly related to PT, and suggest that types K1 and K2 have distinct nucleotide sequences throughout the HCV genome.

To isolate an HCV cDNA fragment from plasma samples negative for the 401 bp fragment by RT-PCR, several primers were designed and tested. A combination of S' and AS' primers (shown in Table 1) produced a 225 bp fragment with higher frequency than that of the 401 bp fragment produced using the S and AS primers, which is particularly useful for the detection of the HCV genome in hepatoma and post-transfusion patients. The HCV genome was detected in 10 samples negative for the 401 bp fragment by RT-PCR, and by restriction enzyme digestion or sequencing; the fragments tested are members of the type K1 group. Although S' and AS' primers facilitate PCR diagnosis of non-A, non-B hepatitis liver disease, the amplified 225 bp fragments are too small for RFLP analysis. These results suggest that there are two main HCV types in Japan, HCV-K1 and HCV-K2.

HCV-PT has been detected frequently in blood from haemophiliacs who have received clotting factors (Hijikata *et al.*, 1990; Ulrich *et al.*, 1990). PT-1 was also isolated from a Japanese haemophiliac, and its nucleotide and amino acid sequences are 95% and 91% identical respectively, to those of PT in a 340 bp portion of the NS5 gene, which results in equivalent *AluI*- and *Sau96I*-generated RFLPs. However, their *AccII*-generated RFLPs are different; PT contains two additional *AccII* sites. Recently, several PT strains have been isolated in our laboratory from blood obtained from European non-A, non-B hepatitis patients and their RFLPs are the same as those of PT or PT-1 (data not shown), suggesting that although the nucleotide sequence of the HCV genome may change during infection, each genotype has been conserved as well as that of the hepatitis B viruses (Okamoto *et al.*, 1987).

The clinical course of disease in each HCV-infected patient has been investigated. Our preliminary observations suggest that types K1 and K2 show different responses to interferon (unpublished results). Further basic and clinical studies are in progress.

We thank Dr Stanley M. Hollenberg of the Fred Hutchinson Cancer Research Center (Seattle, Wash., U.S.A.) for critical reading. This work was supported in part by a grant from the Foundation for Promotion of Cancer Research backed by the Japan Shipbuilding Industry Foundation.

References

- ARGOS, P. (1988). A sequence motif in many polymerases. *Nucleic Acids Research* **16**, 9909-9916.
- CHOMSZYNSKI, P. & SACCHI, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**, 156-159.
- CHOO, Q.-L., WEINER, A. J., OVERBY, L. R., BRADLEY, K. W. & HOUGHTON, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**, 359-362.
- ENOMOTO, N., TAKADA, A., NAKAO, T. & DATE, T. (1990). There are two major types of hepatitis C virus in Japan. *Biochemical and Biophysical Research Communications* **170**, 1021-1025.
- ENOMOTO, N., TAKADA, N., TAKASE, S., TAKADA, A. & DATE, T. (1991). Hepatitis C virus RNA genome in plasma of patients with non-A, non-B hepatitis. *Gastroenterologia Japonica* **26**, 42-46.
- GYLLENSTEN, U. B. & ERLICH, H. A. (1988). Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 7652-7656.
- HIJIKATA, M., KATO, N., MORI, S., OOTSUYAMA, S., NAKAGAWA, M., SUGIMURA, T., OHKOSHI, S., KOJIMA, H., MEGURO, T., TAKI, M., TAKAYAMA, S., YAMADA, K. & SHIMOTOHNO, K. (1990). Frequent detection of hepatitis C virus US strain in Japanese hemophiliacs. *Japanese Journal of Cancer Research* **81**, 1195-1197.
- HOUGHTON, M., CHOO, Q.-L. & KUO, G. (1988). NANB diagnostics and vaccines; *European Patent Application* 88, 310, 9225.5 and Published **318**, 216.
- KANAI, K., IWATA, K., NAKAKO, K., KAKO, M. & OKAMOTO, H. (1990). Suppression of hepatitis C virus RNA by interferon- α . *Lancet* **i**, 245.
- KANEKO, S., UNOURA, M., KOBAYASHI, K., KUNO, K., MURAKAMI, S. & HATTORI, N. (1990). Detection of serum hepatitis C virus RNA. *Lancet* **i**, 976.
- KATO, N., OHKOSHI, S. & SHIMOTOHNO, K. (1989). Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variations from the original isolate in the U.S.A. *Proceedings of the Japan Academy* **65**, 219-223.
- KATO, N., HIJIKATA, M., OOTSUYAMA, Y., NAKAGAWA, M., OHNISHI, S., SUGIMURA, T. & SHIMOTOHNO, K. (1990). Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 9524-9528.
- KUBO, Y., TAKEUCHI, K., BOONMAR, S., KATAYAMA, T., CHOO, Q.-L., KUO, G., WEINER, A. J., BRADLEY, D. W., HOUGHTON, M., SAITO, I. & MIYAMURA, T. (1989). A cDNA fragment of hepatitis C virus isolated from Japan. *Nucleic Acids Research* **17**, 10367-10372.
- KUO, G., CHOO, Q.-L., ALTER, H. J., GITNICK, G. L., REDEKER, A. G., PURCELL, R. H., MIYAMURA, T., DIENSTAG, J. L., ALTER, M. J., STEVENS, C. E., TEGTMER, G. E., BONINO, F., COLOMBO, M., LEE, W.-S., KIO, C., BERGER, K., SHUSTER, J. R., OVERBY, L. R., BRADLEY, D. W. & HOUGHTON, M. (1989). An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* **244**, 362-364.
- MAENO, M., KAMINAKA, K., SUGIMOTO, H., ESUMI, M., HAYASHI, N., KOMATSU, K., ABE, K., SEKIGUCHI, S., YANO, M., MIZUNO, K. & SHIKATA, T. (1990). A cDNA clone closely associated with non-A, non-B hepatitis. *Nucleic Acids Research* **18**, 2685-2689.
- OKAMOTO, H., IMAI, M., TSUDA, F., MIYAKAWA, Y. & MAYUMI, M. (1987). Point mutation in the S gene of hepatitis B virus for d/y or w/r subtypic change in two blood donors carrying a surface antigen of compound subtypes adyt or adwt. *Journal of Virology* **61**, 3030-3034.
- OKAMOTO, H., OKADA, S., SUGIYAMA, Y., YOTSUMOTO, S., TANAKA, T., YOSHIZAWA, H., TSUDA, F., MIYAKAWA, Y. & MAYUMI, M. (1990a). The 5'-terminal sequence of the hepatitis C virus genome. *Japanese Journal of Experimental Medicine* **60**, 167-177.
- OKAMOTO, H., OKADA, S., SUGIYAMA, Y., TANAKA, T., SUGAI, Y., AKAHANE, Y., MACHIDA, A., MISHIRO, S., YOSHIZAWA, H., MIYAKAWA, Y. & MAYUMI, M. (1990b). Detection of hepatitis C virus RNA by two-stage polymerase chain reaction with two pairs of primers deduced from the 5'-noncoding region. *Japanese Journal of Experimental Medicine* **60**, 215-222.
- PLETNEV, A. G., YAMSHCHIKOV, V. F. & BLINOV, V. M. (1990). Nucleotide sequence of the genome and complete amino acid sequence of the polyprotein of tick-borne encephalitis virus. *Virology* **174**, 250-263.

- SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.
- TAKAMIZAWA, A., MORI, C., FUKU, I., MANABE, S., MURAKAMI, S., FUJITA, J., ONISHI, E., ANDOH, T., YOSHIDA, I. & OKAYAMA, H. (1991). Structure and organization of the hepatitis C virus genome isolated from human carriers. *Journal of Virology* **65**, 1105–1113.
- TAKEUCHI, K., KUBO, Y., WATANABE, Y., KATAYAMA, T., CHOO, Q.-L., KUO, G., HOUGHTON, M., SAITO, I. & MIYAMURA, T. (1990*a*). Nucleotide sequence of core and envelope genes of the hepatitis C virus genome derived directly from human healthy carriers. *Nucleic Acids Research* **18**, 4626.
- TAKEUCHI, K., BOONMAR, S., KUBO, Y., KATAYAMA, T., HARADA, H., OHBAYASHI, A., CHOO, Q.-L., KUO, G., HOUGHTON, M., SAITO, I. & MIYAMURA, T. (1990*b*). Hepatitis C viral cDNA clones isolated from a healthy carrier donor implicated in post-transfusion non-A, non-B hepatitis. *Gene* **91**, 287–291.
- TAKEUCHI, K., KUBO, Y., BOONMAR, S., WATANABE, Y., KATAYAMA, T., CHOO, Q.-L., KUO, G., HOUGHTON, M., SAITO, I. & MIYAMURA, T. (1990*c*). The putative nucleocapsid and envelope protein genes of hepatitis C virus determined by comparison of the nucleotide sequences of two isolates derived from an experimentally infected chimpanzee and healthy human carriers. *Journal of General Virology* **71**, 3027–3033.
- ULRICH, P. P., ROMEO, J. M., LANE, P. K., KELLY, I., DANIEL, L. J. & VYAS, G. N. (1990). Detection, semiquantitation, and genetic variation in hepatitis C virus sequences amplified from the plasma of blood donors with elevated alanine aminotransferase. *Journal of Clinical Investigation* **86**, 1609–1614.

(Received 19 March 1991; Accepted 22 May 1991)