

Genetic analysis of the hepatitis C virus (HCV) genome from HCV-infected human T cells

K. Sugiyama, N. Kato, T. Mizutani, M. Ikeda, T. Tanaka and K. Shimotohno

Virology Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104, Japan

We recently established a cell culture system for the replication of hepatitis C virus (HCV) by using a human T cell leukaemia virus type I-infected cell line, MT-2, and showed that the quasi-species of the hypervariable region 1 observed in the original inoculum became homogeneous in MT-2 cells 10 days after inoculation of HCV. In this study, we obtained HCV cDNA clones covering the whole viral genome by RT-nested PCR using RNA from HCV-infected cloned MT-2C cells, which support viral replication more efficiently, at 12 days after inoculation. A total of 41 distinct HCV cDNA clones covering almost the whole viral genome were also isolated from a cDNA library derived from the original inoculum. Molecular evolutionary analyses comparing the sequences of the HCV clones ob-

tained from both sources revealed that the HCV populations became homogeneous in more than half of the compared regions. This finding suggests that limited HCV populations are able to replicate in MT-2C cells. In addition, we isolated cDNA clones containing a 3' X-tail sequence, which was recently identified as a bona fide 3' terminus of the HCV genome, in the HCV-infected MT-2C cells and confirmed that the nucleotide sequence of the 3' X-tail was highly conserved, suggesting its implication in HCV replication. Finally, on the basis of the sequences of HCV cDNA clones obtained from HCV-infected MT-2C cells, we determined the entire nucleotide sequence of the HCV genome containing the 3' X-tail as a candidate for an infectious HCV molecular clone.

Introduction

Hepatitis C virus (HCV) is a major causative agent of chronic hepatitis and frequently gives rise to liver cirrhosis and hepatocellular carcinoma (Ohkoshi *et al.*, 1990; Saito *et al.*, 1990). The viral genome structure and molecular mechanisms of viral protein processing have been elucidated (reviewed by Shimotohno, 1995). However, the biological mechanisms of HCV infection and replication in the target cells are still equivocal. So far, there is no effective treatment for HCV infection other than interferon administration.

HCV is considered to belong to the *Flaviviridae* (Kato *et al.*, 1991; Miller & Purcell, 1990), which have a positive-stranded RNA genome of about 9.5 kb including a large open reading frame (ORF) encoding a polyprotein precursor of about 3000

amino acids. Comprehensive structural analysis of HCV genomes has revealed the existence of at least 9 genotypes and more than 30 subgenotypes throughout the world (reviewed by Bukh *et al.*, 1995; Simmonds, 1995). HCV genotypes show diversities of around 30% in the nucleotide sequences of their whole genomes. Moreover, even in serum from an individual patient, there exist plural viral species harbouring various mutations in their nucleotide sequence, which are called quasi-species. In particular, the hypervariable region 1 (HVR1) located in the N terminus of the envelope 2 (E2) protein shows a marked quasi-species nature (Kato *et al.*, 1992a; Sekiya *et al.*, 1994; Saito *et al.*, 1996). The complexity of the HVR1 quasi-species is thought to play an important role in the escape of HCV from the immunosurveillance system of the host, and in the mechanism of chronic HCV infection (Kato *et al.*, 1993, 1994; Weiner *et al.*, 1992).

Recently, we developed a cell culture system which can support HCV replication by using a human T cell leukaemia virus type I (HTLV-I)-infected T cell line, MT-2, and showed that an HCV population with a limited HVR1 sequence became predominant in the cultured cells, despite the complicated

Author for correspondence: Nobuyuki Kato.

Fax +81 3 3543 2181. e-mail nkato@ncc.go.jp

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under accession number D85516.

quasi-species of the HVR1 in the primary inoculum (Kato *et al.*, 1995). To determine whether the same phenomenon occurs throughout the HCV genome, we carried out RT-nested PCR on the whole viral genome from HCV-infected MT-2C cells, a clone from MT-2 cells which was more susceptible to HCV (Mizutani *et al.*, 1995, 1996). We determined the nucleotide sequences of these PCR products and compared them with those of clones isolated from a cDNA library derived from the inoculum. These comparisons revealed that convergence of HCV populations, as observed in the HVR1, occurred in the MT-2C cell culture after inoculation. Moreover, we also isolated cDNA clones containing a 3' X-tail sequence, a novel nucleotide sequence recently discovered to be a bona fide 3' terminus of the viral genome (Tanaka *et al.*, 1995, 1996), from HCV-infected MT-2C cells, suggesting its implication in HCV replication in cultured T cells.

Methods

■ **Inoculum.** The inoculum used in this study was serum 1B-1 [identical to serum #1090 in a previous report by Kato *et al.* (1995)] supplied by The Japan Red Cross Central Blood Center, Tokyo. This serum was found to contain about 10^7 HCV genomes (HCV genotype 1b) per ml.

■ **Construction and screening of a cDNA library from serum 1B-1.** A 40 ml sample of serum 1B-1 was centrifuged at 25 000 r.p.m. for 5 h, and then RNA was prepared from the pellet using ISOGEN-LS extraction kits (Nippon Gene, Japan). A cDNA library of RNA from serum 1B-1 was prepared in λ gt10 using several primers which were designed on the basis of the nucleotide sequences of the conserved region of the HCV genome (Tanaka *et al.*, 1992), according to the manufacturer's protocol for cDNA synthesis and the λ gt10 cloning system (Amersham). The cDNA library was screened by using cDNA fragments of the HCV-J genome (Kato *et al.*, 1990b) as probes.

■ **Subcloning of cDNA clones from the cDNA library.** The cDNA inserts in λ gt10 were subcloned into the plasmid vector pTZ19R and used for sequence analysis. Nucleotide sequencing was done with an AutoRead sequencing kit (Pharmacia). Nucleotide positions used in this report correspond to those of HCV-JT (Tanaka *et al.*, 1992). Nucleotide positions within the 3' X-tail sequence correspond to those in a previous report by Tanaka *et al.* (1995).

■ **Cell culture and virus inoculation.** Cloned MT-2C cells, which were more susceptible to HCV than uncloned MT-2 cells, were cultured in RPMI-1640 supplemented with 10% foetal bovine serum (Mizutani *et al.*, 1995). One ml of cell suspension (1×10^6 cells) was mixed with 0.5 ml of undiluted serum 1B-1, incubated for 20 h at 37 °C, and then treated as described previously (Kato *et al.*, 1995). In a previous study, we demonstrated that HCV RNA synthesis in infected MT-2C cells increased after inoculation, reached a maximum level at 4 days post-inoculation (p.i.), and maintained this level until at least 11 days p.i. (Mizutani *et al.*, 1996). Negative-stranded HCV RNA was also detected in the infected cells by two different methods with strand specificity (Mizutani *et al.*, 1996). These results suggest that HCV replicates and multiplies in MT-2C cells.

■ **RT-nested PCR of RNA from HCV-infected MT-2C cells.** At 12 days p.i., RNAs were extracted from MT-2C cells as described previously (Kato *et al.*, 1995). One-fifth of the RNAs were used for RT-nested PCR

as described previously (Kato *et al.*, 1992b). In short, after cDNA synthesis with reverse transcriptase SuperScript II (GIBCO BRL), the reaction mixture was boiled at 98 °C for 10 min to inactivate the reverse transcriptase, and then synthesized cDNA was amplified, first by Taq DNA polymerase (Sawady Technology, Japan) with an outer primer set in an automatic thermocycler (Perkin Elmer Cetus). For the second amplification, 2 μ l of the first reaction mixture was removed and further amplified with an inner primer set. All primers were designed mainly on the basis of the nucleotide sequences of the conserved region of the HCV genome (Tanaka *et al.*, 1992). In this way, 17 regions (regions B to R, closed boxes in Fig. 1) were amplified.

Region S, including the conventional 3' non-coding region, the poly(U) stretch and the 5' part of the 3' X-tail (Fig. 1), was reverse-transcribed with primer R4 and amplified with two sets of primers, 9348/R5 (in the first reaction) and 9363Xh/R6Xh (in the second reaction). In addition region T, located within the 3' X-tail region (Fig. 1), was reverse-transcribed with primer RP2 and amplified with two sets of primers, U6/RP2 (in the first reaction) and U5/R7 (in the second reaction). The nucleotide sequences of these primers, except for R4 (5' GGCCAGTATCAGCACTCTCT 3', nucleotides 63'-82') and R7 (5' TCTGCAGAGAGGCCAGTATC 3', 73'-92'), were described in a previous report (Tanaka *et al.*, 1996).

For comparison, regions S and T of the HCV genome from serum 1B-1 were also amplified by the method described above.

■ **Determination of the 5' terminus of the HCV genome by the primer extension-anchored PCR method.** The 5' extremity of the HCV genome from the HCV-infected MT-2C cells (region A in Fig. 1) was determined by the primer extension-anchored PCR method as follows. Briefly, cDNA was synthesized with primer 362R (5' GTAC-CACAAGGCCTTTCGCG 3', 270-289), and then a Centricon-30 microconcentrator (Amicon) was used to remove excess primers and dNTPs. The dC (or dA) tailing reaction mixture (20 μ l) containing 0.1 M potassium cacodylate (pH 7.2), 2 mM CoCl_2 , 0.2 mM DTT, 0.2 mM dCTP (or dATP), synthesized cDNA and 15 U of terminal deoxynucleotidyl transferase (GIBCO BRL) was incubated at 37 °C for 1 h. Then, 1 μ l of the reaction mixture was subjected to nested PCR. The primers used were as follows: for the first PCR reaction, 197R (antisense, 5' CTTTCGCGACCCAACACTAC 3', 258-277) and 363 [sense for dC-tailing, 5' GACTCGAGTGGATCCATCGA(G)₁₇ 3'] or 364 [sense for dA-tailing, 5' GACTCGAGTGGATCCATCGA(T)₂₅ 3']; for the second PCR reaction, AdpBam (sense, 5' GACTCGAGTGGATCCATCGA 3') containing a *Bam*HI site (underlined) and 361RE (antisense, 5' GACTGTACAGAATTCCCAAATCTCCAGGCATTGAG 3', 211-230) containing an *Eco*RI site (underlined).

■ **Cloning of PCR products.** The PCR products were digested by restriction enzymes whose sites were present in the primers, or blunted by Klenow fragment, phosphorylated with T4 polynucleotide kinase as described previously (Kato *et al.*, 1990a), and then subcloned into a pTZ19R or a Bluescript II KS(+) plasmid vector for sequence analysis.

■ **Molecular evolutionary analysis.** Nucleotide or deduced amino acid sequences of the clones isolated from the serum 1B-1 cDNA library and the clones obtained by RT-nested PCRs from MT-2C cells inoculated with serum 1B-1 were analysed by the unweighted pair-grouping method with arithmetic mean (Nei & Gojobori, 1986) with the program GENETYX-MAC (Software Development, Japan). In addition, six reported HCV genomes belonging to genotype 1b-1 (HCV-J, Kato *et al.*, 1990b; HCV-JT, Tanaka *et al.*, 1992; HCV-BK, Takamizawa *et al.*, 1991; HCV-T, Chen *et al.*, 1992; HCV-JK1, Honda *et al.*, 1992; and HCV-J33, GenBank accession number D14484) were used as references. Seventeen

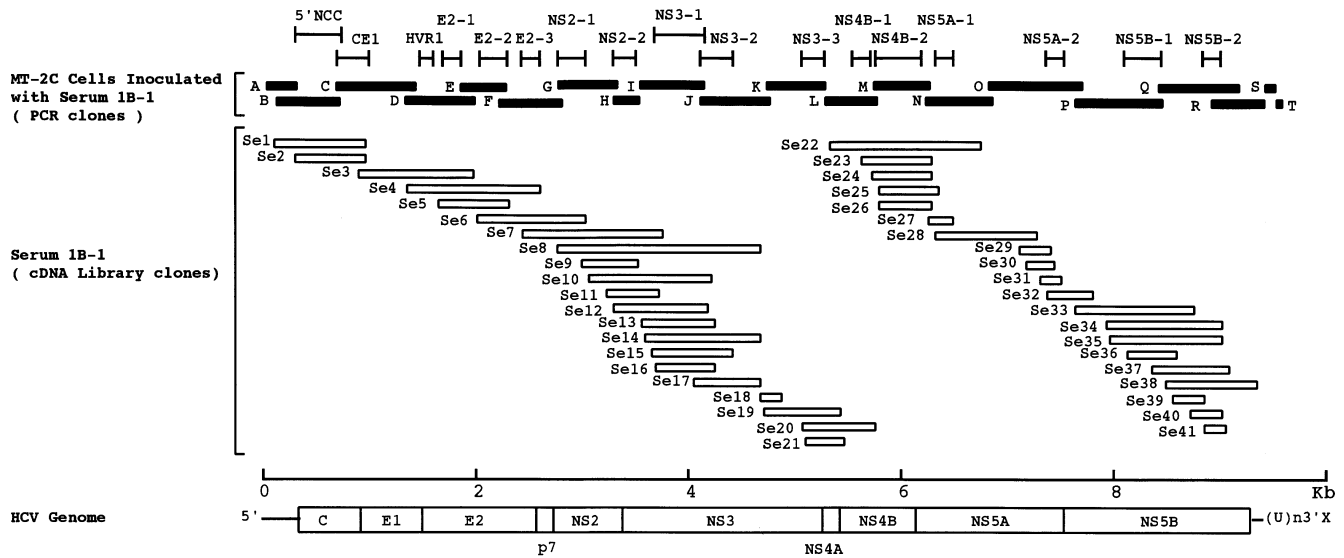


Fig. 1. Map of clones isolated from the cDNA library of serum 1B-1 and RT-nested PCR products from cultured MT-2C cells 12 days after inoculation with serum 1B-1. The general structure of the full-length HCV genome is illustrated at the bottom of the figure. Forty-one clones isolated from the cDNA library are presented as open boxes and designated Se1 to Se41. The RT-nested PCR products from the cultured MT-2C cells are presented as filled boxes and designated by letters A to T. Nucleotide positions of the amplified regions, excluding PCR primers, are as follows: A (1–210), B (102–696), C (640–1421), D (1322–1960), E (1850–2276), F (2204–2771), G (2762–3314), H (3281–3488), I (3446–4136), J (4102–4752), K (4730–5272), L (5267–5759), M (5735–6245), N (6245–6848), O (6827–7688), P (7646–8450), Q (8429–9170), R (8877–9382), S (9383–28'), T (16'–72'). Numbers of nucleotide positions in parentheses correspond to those of HCV-JT (Tanaka *et al.*, 1992). Numbers with a prime correspond to those in the published 3' X-tail sequence (Tanaka *et al.*, 1995). The 17 regions compared by molecular evolutionary analysis are indicated by thin bars at the top of the figure. Nucleotide positions of each analysed region are as follows: 5'NCC (273–696), CE1 (640–915), HVR1 (1493–1571), E2-1 (1600–1853), E2-2 (1953–2275), E2-3 (2389–2566), NS2-1 (2762–2972), NS2-2 (3281–3473), NS3-1 (3646–4136), NS3-2 (4102–4346), NS3-3 (4102–4346), NS4B-1 (5563–5709), NS4B-2 (5735–6238), NS5A-1 (6275–6426), NS5A-2 (7300–7447), NS5B-1 (8057–8450), NS5B-2 (8805–8936).

regions which contain more than 120 nucleotides each were chosen for the analysis and referred to as 5'NCC, CE1, HVR1, E2-1, E2-2, E2-3, NS2-1, NS2-2, NS3-1, NS3-2, NS3-3, NS4B-1, NS4B-2, NS5A-1, NS5A-2, NS5B-1 and NS5B-2 (upper part of Fig. 1).

Results

Isolation of cDNA clones covering whole HCV genome

We obtained 41 HCV cDNA clones from the cDNA library constructed with serum 1B-1 used as inoculum. Sequence analyses revealed that these cDNA clones (Se1 to Se41, open boxes in Fig. 1) spanned almost the entire HCV genome [nucleotide positions 52–9265 of HCV-JT (Tanaka *et al.*, 1992)]. In addition, we successfully amplified the 20 regions (A to T, closed boxes in Fig. 1) covering the whole HCV genome by RT-nested PCR using RNA from HCV-infected MT-2C cells at 12 days p.i. Basically, at least three clones from each PCR product were sequenced to determine the status of the quasi-species. For further characterization, we sequenced 11, 18, 5 and 18 clones from regions A, D, S and T, respectively.

Molecular evolutionary analysis of the deduced amino acid sequences of the HVR1

Since serum 1B-1 showed complicated quasi-species of HVR1 populations consisting of 4 subgroups and 22 distinct

species (Kato *et al.*, 1996), we first analysed the HVR1 region. Fig. 2(a) shows the phylogenetic tree of deduced amino acid sequences of the HVR1 species obtained from serum 1B-1 (42 clones, classified as I-1 to IV-9), and those from the HCV-infected MT-2C cells at 12 days p.i. (18 clones, D1–D18). As a consequence, only five HVR1 species [II-1 (9 clones), III-1 (4 clones), I-1 (3 clones), II-6 (1 clone) and II-7 (1 clone)] were obtained from HCV-infected MT-2C cells, although 23 distinct HVR1 species were obtained from serum 1B-1. In particular, species II-1 became predominant in the MT-2C cells with a frequency of 50%, although its frequency was only 13% in the inoculum. Moreover, HVR1 species II-6 and II-7, which differed from species II-1 by only one amino acid and were not isolated from the inoculum, were also obtained from the MT-2C cells. These results indicate that it is mainly subgroup II which is selected in the MT-2C cell culture after virus inoculation.

Molecular evolutionary analysis of the deduced amino acid sequences of the E2-1 region

To determine whether the selection of HCV species, as observed in the HVR1, occurs in other regions, we analysed the E2-1 region (84 aa), which is located just downstream from the HVR1. We made a phylogenetic tree for this region from three cDNA library clones (Se3, Se4 and Se5) and 18 PCR

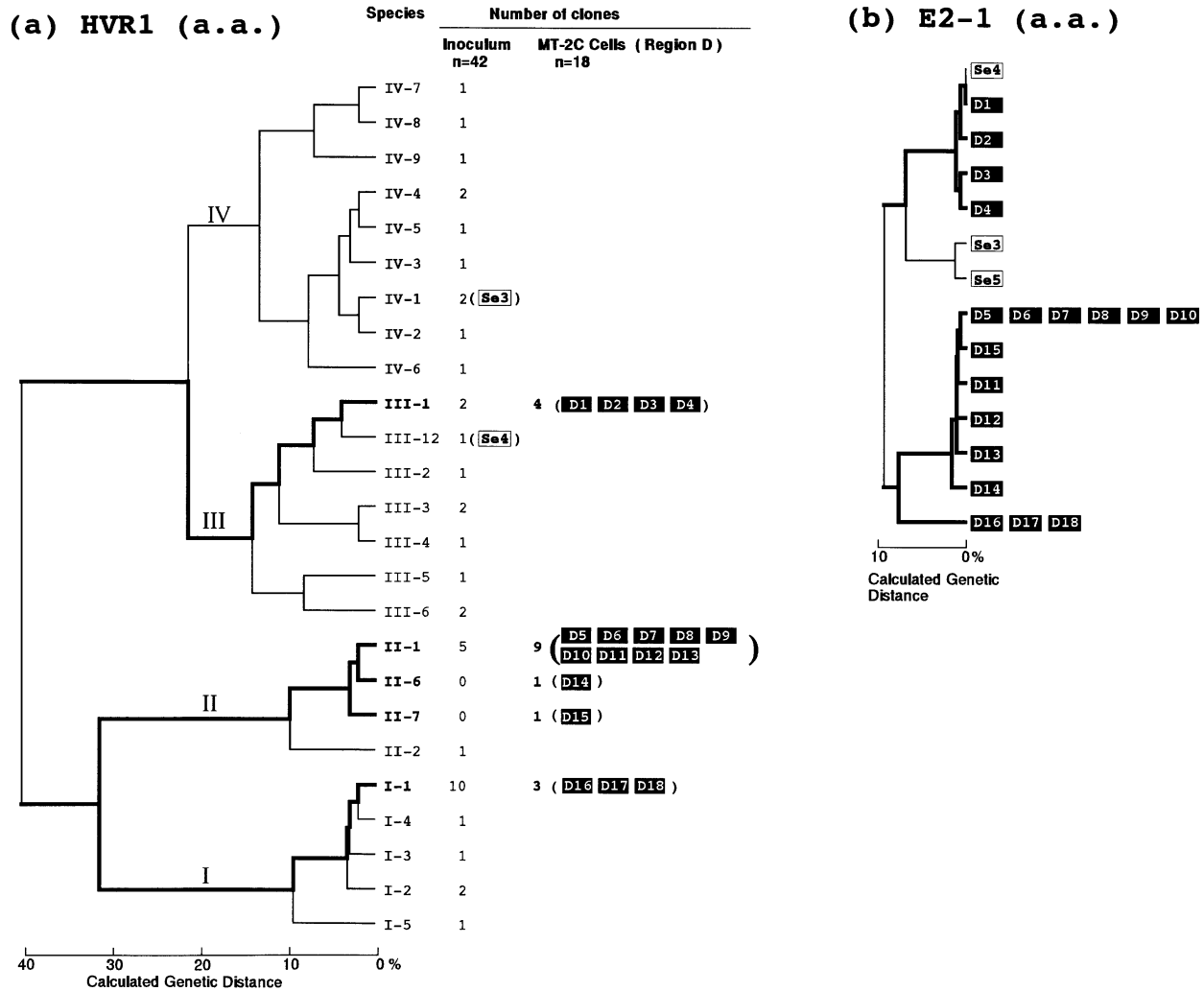


Fig. 2. Molecular evolutionary analyses of amino acid sequences in the HVR1 and E2-1 regions within the HCV E2 protein from serum 1B-1 and MT-2C cells 12 days after inoculation with serum 1B-1. (a) A phylogenetic tree was depicted on the basis of amino acid sequences of the HVR1 from serum 1B-1 ($n = 40$) (Kato *et al.*, 1996) and the PCR clones in region D ($n = 18$; D1 to D18 in filled boxes) from the MT-2C cells. Two available clones (Se3 and Se4) from the cDNA library were added to the data of serum 1B-1. Main branches are referred to as I, II, III and IV. Species are assigned a branch number followed by an individual number (for example, IV-7). Branches including the PCR clones in region D are shown with bold lines. (b) A phylogenetic tree is depicted on the basis of amino acid sequences of three cDNA library clones of the inoculum (Se3, Se4 and Se5 in open boxes) and 18 PCR clones (D1 to D18 in filled boxes). Branches which include the PCR clones are shown with bold lines as in (a).

clones (D1 to D18). As shown in Fig. 2(b), a phylogenetic tree for E2-1 had almost the same branching as that for the HVR1, indicating that the sequence relationships in the HVR1 were equivalent to those in the E2-1 region.

Molecular evolutionary analysis of the nucleotide sequences of 16 regions other than the HVR1

We further analysed 16 regions other than the HVR1 (from 5'NCC to NS5B-2, Fig. 3). For each region, phylogenetic trees were constructed by using nucleotide sequence data of the cDNA library clones and of PCR clones obtained from the HCV-infected MT-2C cells (open and closed boxes, respec-

tively, in Fig. 3). To determine their relative genetic distances, the nucleotide sequences of six previously reported HCV genomes (HCV-J, -JT, -BK, -T, -JK1 and -J33) belonging to genotype 1b were added to the analysis for each phylogenetic tree, as shown in Fig. 3. Focusing on the PCR clones and the cDNA library clones in the E2-1 region reveals that the phylogenetic tree based on the nucleotide sequence data is almost the same as that obtained from the amino acid sequence data in Fig. 2(b).

Phylogenetic trees in the remaining 15 regions revealed that sequence diversities among PCR clones and cDNA library clones were less than those among the six reported HCV strains, although only Se22 showed a rather greater genetic

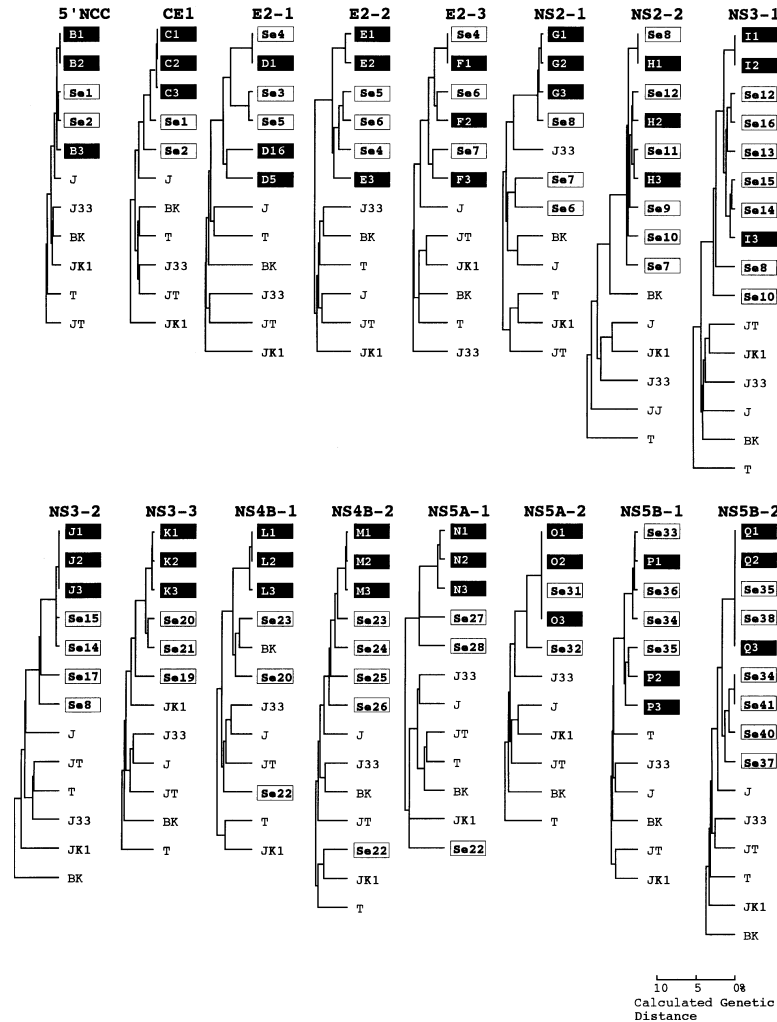


Fig. 3. Molecular evolutionary analyses of nucleotide sequences in 16 regions throughout the HCV genome from the serum 1B-1 and the MT-2C cells 12 days after inoculation with serum 1B-1. Nucleotide positions of the analysed regions are described in the legend to Fig. 1. Phylogenetic trees are depicted on the basis of nucleotide sequences of cDNA library clones (open boxes), which were available in each region, and three PCR clones in each region (filled boxes). For comparisons, nucleotide sequence data of six representative HCV isolates belonging to the HCV-1b genotype are added to the analyses: J, HCV-J; JT, HCV-JT; BK, HCV-BK; T, HCV-T; JK1, HCV-JK1; and J33, HCV-J33.

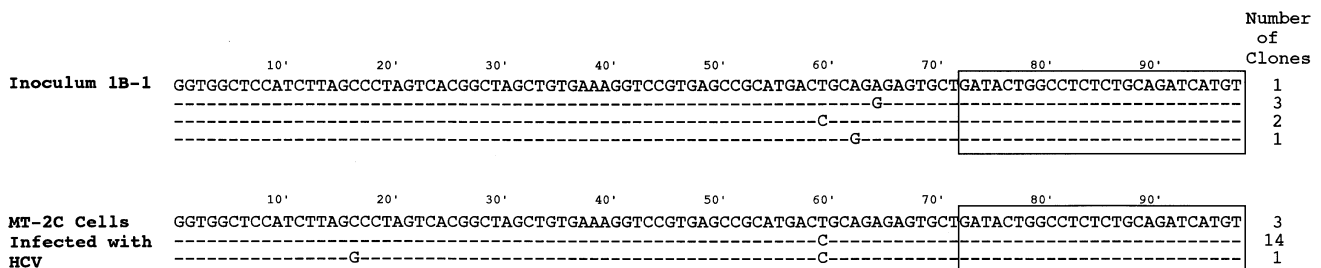


Fig. 4. Nucleotide sequences of the 3' X-tail cDNAs from serum 1B-1 and the MT-2C cells 12 days after inoculation with serum 1B-1. The nucleotide sequence in the top line is identical to the consensus 3' X-tail sequence published previously (Tanaka *et al.*, 1995). Numbers with a prime correspond to the nucleotide position in the published sequence. Nucleotides identical to those in the top line are denoted by dashes. Boxed areas indicate PCR primer sequences.

distance. In particular, in 10 regions (5'NCC, CE1, NS2-1, NS3-2, NS3-3, NS4B-1, NS4B-2, NS5A-1, NS5A-2 and NS5B-2), three PCR clones from MT-2C cells showed almost the same nucleotide sequences, despite an apparent sequence heterogeneity among the cDNA library clones from serum 1B-1. These results suggest that a limited HCV population became predominant in the HCV-infected MT-2C cells, although the PCR clones in the E2-2, E2-3, NS2-2, NS3-1 and NS5B-1

regions were still as heterogeneous as the cDNA library clones.

Analysis of the 3' X-tail

The recently identified 3'-terminal sequence of the HCV genome, 3' X-tail (Tanaka *et al.*, 1995, 1996), was successfully amplified by using RNAs from both serum 1B-1 and MT-2C cells inoculated with serum 1B-1. As shown in Fig. 4, the 3' X-

tail sequence (98 nucleotides) was relatively conserved even in serum 1B-1, as observed previously (Tanaka *et al.*, 1995, 1996). The 3' X-tail sequences from HCV-infected MT-2C cells were also highly conserved, although minor heterogeneity was observed. A clone with a substitution (T to C) at nucleotide position 60' was found in 15 out of 18 clones in the HCV-infected MT-2C cells, although in only 2 out of 7 clones in serum 1B-1.

The poly(U) stretch, which is located just upstream of the 3' X-tail, was also amplified and characterized. As observed in our previous studies (Tanaka *et al.*, 1995, 1996), the poly(U) stretch from this inoculum also showed marked heterogeneity in length, which varied from 47 to 71 nucleotides (average of 60) (data not shown). However, interestingly, the poly(U) stretch from the HCV-infected MT-2C cells was apparently shorter than that from serum 1B-1, and the range of the lengths was more restricted, from 29 to 32 nucleotides (average of 30), suggesting that a shorter poly(U) stretch might affect the replication of HCV in MT-2C cells. Within the poly(U) stretch region, as observed previously (Tanaka *et al.*, 1995, 1996), a variable distribution of C, CC or CCC residues was also observed (data not shown).

Entire nucleotide sequence of the HCV genome from the HCV-infected MT-2C cells

To identify the HCV genome replicating in MT-2C cells, we determined the consensus nucleotide sequence from the PCR clones in each region, except for region D, in which we chose D5 itself as a representative sequence since it is a major HVR1 sequence (II-1). The entire HCV genome was 9535 nucleotides long (including the primer sequence in the extremity of the 3' X-tail), consisting of 341 nucleotides of the 5' non-coding region, 9030 nucleotides of the ORF and 164 nucleotides of the 3' non-coding region including the poly(U) stretch and the 3' X-tail. There was no deletion, insertion or stop codon in the ORF comparable with that of HCV-JT (Tanaka *et al.*, 1992). The sequence was designated HCV-JS and deposited in the nucleotide sequence databases under accession number D85516.

Discussion

To date there have been several studies in which HCV infection of cultured cells has been demonstrated (Cribier *et al.*, 1995; Lanford *et al.*, 1994; Shimizu *et al.*, 1993; Zibert *et al.*, 1995). We have also previously demonstrated replication and multiplication of HCV in cultured MT-2 cells (a T cell line infected with HTLV-1) (Kato *et al.*, 1995), cloned MT-2C cells (Mizutani *et al.*, 1995, 1996) and PH5CH cells (Kato *et al.*, 1996), a non-neoplastic hepatocyte line immortalized by simian virus 40 large T antigen (Noguchi & Hirohashi, 1996). Intriguingly, in these systems, HVR1 populations became homogeneous during culture despite the presence of complicated quasi-species in the inoculum (Kato *et al.*, 1995, 1996).

A similar phenomenon was observed in the HCV-infected T cell line HPB-Ma (Hijikata *et al.*, 1995). In the present study, we demonstrated homogeneity in more than half of the compared regions of the HCV genome obtained from HCV-infected MT-2C cells. We also determined the nucleotide sequence of the entire HCV genome (HCV-JS) and this sequence may represent that of the infectious molecular clone.

The predominant HVR1 species, II-1, obtained from the MT-2C cells at 12 days p.i. was different from the predominant species, I-2, observed in uncloned MT-2 cells at 10 days p.i. (Kato *et al.*, 1995). One possible explanation for this conflict is that HVR1 species II-1 was amplified by using a newly designed primer (Kato *et al.*, 1996) which was not used in the analysis of uncloned MT-2 cells. This primer was designed to have greater homology with the nucleotide sequences of most HCV strains registered to date in the DDBJ/EMBL/GenBank databases. HVR1 species I-2 could be efficiently amplified by using this newly designed primer. However, since HVR1 species I-2 was not isolated from cloned MT-2C cells, it is unlikely that this discrepancy was due to the difference in the primers used. The other possible explanation is a difference in the proliferation of an HCV clone in uncloned and cloned MT-2C cells. This difference might lead to the accumulation of different HVR1 species in those cells.

By comparing cDNA clones covering the HVR1 and the adjacent E2-1 region, we demonstrated that the relationship of amino acid sequences within these regions was equivalent. This finding supports the presumption that serum 1B-1 contains four major HCV subgroups (I to IV) (Kato *et al.*, 1996). Moreover, all four subgroups are located in a separate branch from the other six HCV isolates (Fig. 3), suggesting that they may have evolved recently from a single HCV ancestor. Evolutionary analysis of cDNA clones obtained in 17 regions throughout the whole viral genome revealed that, in 10 regions, HCV populations became homogeneous during the 12 day culture period after virus inoculation. However, we obtained cDNA clones showing heterogeneity of nucleotide sequences in other regions including the HVR1. For complete homogeneity of HCV populations, a prolonged culture period of more than 13 days after virus inoculation may be required. A similar result was obtained in another HCV-replicating cell culture system using a human T cell line, HPB-Ma cells (Nakajima *et al.*, 1996). Taken together, these results suggest that only a limited HCV population in the HCV inoculum is able to adapt to replication in the restricted surroundings.

During analysis of the 5' terminus of the HCV genome (region A in Fig. 1) by the primer extension-anchored PCR method, we found an additional U residue at the extremity of the 5'NC region in only 1 out of 11 PCR clones. Although such a U residue at the 5' extremity of the HCV genome was also found in the analysis of the HCV-J33 genome (GenBank accession number D14484), its implied role in viral translation or replication needs to be determined. Low efficiency of reverse-transcriptase activity at the very extremity of a template

may account for the low frequency of the extra U residue. However, there is also a possibility that it had been accidentally incorporated during the experiments.

In our previous study, the 3' X-tail sequence was detected in the livers and sera of type-C hepatitis patients, and was highly conserved among HCV isolates. In this study, we demonstrated the existence of the 3' X-tail sequence of the HCV genome in an HCV-replicating cultured cell line. Although the nucleotide sequence of the 3' X-tail in serum 1B-1 was relatively conserved, the population of the 3' X-tail sequence in the HCV-infected MT-2C cells became more homogeneous than observed for other regions of the viral genome. Interestingly, in addition, the length of the poly(U) stretch in the HCV-infected MT-2C cells was restricted to between 29 and 32 nucleotides, which is shorter than that (41 to 70 nucleotides) in the original inoculum. In our previous study, we observed that the poly(U) stretch was slightly shortened during PCR amplification (Tanaka *et al.*, 1996). However, it is unlikely that the shortening of the poly(U) stretch occurs simply through experimental procedures such as PCR amplification or an enzymatic sequencing reaction, because the shortening of the poly(U) stretch in the HCV-infected cells was much greater than would be explained by artificial causes. Further analysis using another cell culture system for HCV replication, such as PH5CH hepatocytes (Kato *et al.*, 1996), will be needed to reach a conclusion. In any case, the 3' X-tail and the poly(U) stretch are expected to play an important role in replication of HCV.

We determined the nucleotide sequence of the whole HCV genome, designated HCV-JS, by standardization of nucleotide sequences of PCR cDNA clones obtained from HCV-infected MT-2C cells. Molecular evolutionary analysis of the amino acid sequences of the whole HCV ORF revealed that the genetic distance of HCV-JS from other HCV-1b isolates was just compatible with that among the six previously reported HCV-1b isolates (data not shown). This finding indicates that isolate HCV-JS is not discrete from other HCV-1b isolates. Infectious cDNA clones of flaviviruses or pestiviruses have been obtained mostly from viruses which adapted to cultured cells (Khromykh & Westaway, 1994; Lai *et al.*, 1991; Moormann *et al.*, 1996; Rice *et al.*, 1989; Sumiyoshi *et al.*, 1992). In this context, it is presumed that HCV-JS is very closely related to an infectious HCV genome which is capable of replication in MT-2C cells. It may be possible, with the availability of the sequence information of HCV-JS, to engineer the whole HCV genome as an infectious HCV molecular clone. This may facilitate a better understanding of the precise mechanisms of HCV replication and multiplication, and may lead to the development of an effective treatment for HCV infection.

We thank Ms T. Kobayashi for her helpful assistance. This work was supported by Grants-in-Aid for Cancer Research and for the Second-Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare, and Grants-in-Aid for Scientific Research

from the Ministry of Education, Science and Culture of Japan. M. Ikeda is a recipient of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research, Japan.

References

- Bukh, J., Miller, R. H. & Purcell, R. H. (1995). Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. [Review]. *Seminars in Liver Disease* **15**, 41–63.
- Chen, P. J., Lin, M. H., Tai, K. F., Liu, P. C., Lin, C. J. & Chen D. S. (1992). The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA. *Virology* **188**, 102–113.
- Cribier, B., Schmitt, C., Bingen, A., Kirn, A. & Keller, F. (1995). *In vitro* infection of peripheral blood mononuclear cells by hepatitis C virus. *Journal of General Virology* **76**, 2485–2491.
- Hijikata, M., Mizuno, K., Rikihisa, T., Shimizu, Y. K., Iwamoto, A., Nakajima, N. & Yoshikura, H. (1995). Selective transmission of hepatitis C virus in vivo and in vitro. *Archives of Virology* **140**, 1623–1628.
- Honda, M., Kaneko, S., Unoura, M., Kobayashi, K. & Murakami, S. (1992). Sequence comparisons for a hepatitis C virus genome RNA isolated from a patient with liver cirrhosis. *Gene* **120**, 317–318.
- Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S. & Shimotohno, K. (1990a). Sequence diversity of hepatitis C viral genomes. *Molecular Biology & Medicine* **7**, 495–501.
- Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T. & Shimotohno, K. (1990b). 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, USA* **87**, 9524–9528.
- Kato, N., Hijikata, M., Nakagawa, M., Ootsuyama, Y., Muraiso, K., Ohkoshi, S. & Shimotohno, K. (1991). Molecular structure of the Japanese hepatitis C viral genome. *FEBS Letters* **280**, 325–328.
- Kato, N., Ootsuyama, Y., Ohkoshi, S., Nakazawa, T., Sekiya, H., Hijikata, M. & Shimotohno, K. (1992a). Characterization of hypervariable regions in the putative envelope protein of hepatitis C virus. *Biochemical and Biophysical Research Communications* **189**, 119–127.
- Kato, N., Ootsuyama, Y., Tanaka, T., Nakagawa, M., Nakazawa, T., Muraiso, K., Ohkoshi, S., Hijikata, M. & Shimotohno, K. (1992b). Marked sequence diversity in the putative envelope proteins of hepatitis C viruses. *Virus Research* **22**, 107–123.
- Kato, N., Sekiya, H., Ootsuyama, Y., Nakazawa, T., Hijikata, M., Ohkoshi, S. & Shimotohno, K. (1993). Humoral immune response to hypervariable region I of the putative envelope glycoprotein (gp70) of hepatitis C virus. *Journal of Virology* **67**, 3923–3930.
- Kato, N., Ootsuyama, Y., Sekiya, H., Ohkoshi, S., Nakazawa, T., Hijikata, M. & Shimotohno, K. (1994). Genetic drift in hypervariable region I of the viral genome in persistent hepatitis C virus infection. *Journal of Virology* **68**, 4776–4784.
- Kato, N., Nakazawa, T., Mizutani, T. & Shimotohno, K. (1995). Susceptibility of human T-lymphotropic virus type I infected cell line MT-2 to hepatitis C virus infection. *Biochemical and Biophysical Research Communications* **206**, 863–869.
- Kato, N., Ikeda, M., Mizutani, T., Sugiyama, K., Noguchi, M., Hirohashi, S. & Shimotohno, K. (1996). Replication of hepatitis C virus in cultured non-neoplastic human hepatocytes. *Japanese Journal of Cancer Research* **87**, 787–792.
- Khromykh, A. A. & Westaway, E. G. (1994). Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA. *Journal of Virology* **68**, 4580–4588.

- Lai, C. J., Zhao, B. T., Hori, H. & Bray, M. (1991). Infectious RNA transcribed from stably cloned full-length cDNA of dengue type 4 virus. *Proceedings of the National Academy of Sciences, USA* **88**, 5139–5143.
- Lanford, R. E., Sureau, C., Jacob, J. R., White, R. & Fuerst, T. R. (1994). Demonstration of in vitro infection of chimpanzee hepatocytes with hepatitis C virus using strand-specific RT/PCR. *Virology* **202**, 606–614.
- Miller, R. H. & Purcell, R. H. (1990). Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proceedings of the National Academy of Sciences, USA* **87**, 2057–2061.
- Mizutani, T., Kato, N., Hirota, M., Sugiyama, K., Murakami, A. & Shimotohno, K. (1995). Inhibition of hepatitis C virus replication by antisense oligonucleotide in culture cells. *Biochemical and Biophysical Research Communications* **212**, 906–911.
- Mizutani, T., Kato, N., Saito, S., Ikeda, M., Sugiyama, K. & Shimotohno, K. (1996). Characterization of hepatitis C virus replication in cloned cells obtained from a human T-cell leukemia virus type I infected cell line, MT-2. *Journal of Virology* **70**, 7219–7223.
- Moormann, R. J. M., Genip, H. G. P., Miedema, G. K. W., Hulst, M. M. & Rijn, P. A. (1996). Infectious RNA transcribed from an engineered full-length cDNA template of the genome of a pestivirus. *Journal of Virology* **70**, 763–770.
- Nakajima, N., Hijikata, M., Yoshikura, H. & Shimizu, Y. K. (1996). Characterization of long-term cultures of hepatitis C virus. *Journal of Virology* **70**, 3325–3329.
- Nei, M. & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418–426.
- Noguchi, M. & Hirohashi, S. (1996). Cell lines from non-neoplastic liver and hepatocellular carcinoma tissue from a single patient. *In Vitro Cellular & Developmental Biology Animal* **32**, 135–137.
- Ohkoshi, S., Kojima, H., Tawaraya, H., Miyajima, T., Kamimura, T., Asakura, H., Satoh, A., Hirose, S., Hijikata, M., Kato, N. & Shimotohno, K. (1990). Prevalence of antibody against non-A, non-B hepatitis virus in Japanese patients with hepatocellular carcinoma. *Japanese Journal of Cancer Research* **81**, 550–553.
- Rice, C. M., Grakoui, A., Galler, R. & Chambers, T. J. (1989). Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation. *New Biologist* **1**, 285–296.
- Saito, I., Miyamura, T., Ohbayashi, A., Harada, H., Katayama, T., Kikuchi, S., Watanabe, Y., Koi, S., Onji, M., Ohta, Y., Choo, Q.-L., Houghton, M. & Kuo, G. (1990). Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. *Proceedings of the National Academy of Sciences, USA* **87**, 6547–6549.
- Saito, S., Kato, N., Hijikata, M., Gunji, T., Itabashi, M., Kondo, M., Tanaka, K. & Shimotohno, K. (1996). Comparison of hypervariable regions (HVR1 and HVR2) in positive- and negative-stranded hepatitis C virus RNA in cancerous and non-cancerous liver tissue, peripheral blood mononuclear cells and serum from a patient with hepatocellular carcinoma. *International Journal of Cancer* **67**, 199–203.
- Sekiya, H., Kato, N., Ootsuyama, Y., Nakazawa, T., Yamauchi, K. & Shimotohno, K. (1994). Genetic alterations of the putative envelope proteins encoding region of the hepatitis C virus in the progression to relapsed phase from acute hepatitis: humoral immune response to hypervariable region 1. *International Journal of Cancer* **57**, 664–670.
- Shimizu, Y. K., Purcell, R. H. & Yoshikura, H. (1993). Correlation between the infectivity of hepatitis C virus in vivo and its infectivity in vitro. *Proceedings of the National Academy of Sciences, USA* **90**, 6037–6041.
- Shimotohno, K. (1995). Hepatitis C virus as a causative agent of hepatocellular carcinoma. *Intervirology* **38**, 162–169.
- Simmonds, P. (1995). Variability of hepatitis C virus. [Review]. *Hepatology* **21**, 570–583.
- Sumiyoshi, H., Hoke, C. H. & Trent, D. W. (1992). Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates. *Journal of Virology* **66**, 5425–5431.
- Takamizawa, A., Mori, C., Fuke, 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.
- Tanaka, T., Kato, N., Nakagawa, M., Ootsuyama, Y., Cho, M. J., Nakazawa, T., Hijikata, M., Ishimura, Y. & Shimotohno, K. (1992). Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals. *Virus Research* **23**, 39–53.
- Tanaka, T., Kato, N., Cho, M. J. & Shimotohno, K. (1995). A novel sequence found at the 3' terminus of hepatitis C virus genome. *Biochemical and Biophysical Research Communications* **215**, 744–749.
- Tanaka, T., Kato, N., Cho, M. J., Sugiyama, K. & Shimotohno, K. (1996). Structure of the 3' terminus of the hepatitis C virus genome. *Journal of Virology* **70**, 3307–3312.
- Weiner, A. J., Geysen, H. M., Christopherson, C., Hall, J. E., Mason, T. J., Saracco, G., Bonino, F., Crawford, K., Marion, C. D., Crawford, K. A., Brunetto, M., Bar, P. J., Miyamura, T., McHutchinson, J. & Houghton, M. (1992). Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: potential role in chronic HCV infections. *Proceedings of the National Academy of Sciences, USA* **89**, 3468–3472.
- Zibert, A., Schreier, E. & Roggendorf, M. (1995). Antibodies in human sera specific to hypervariable region 1 of hepatitis C virus can block viral attachment. *Virology* **208**, 653–661.

Received 1 July 1996; Accepted 24 October 1996