Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 3392–3396, April 1991
Medical Sciences

Nucleotide sequence and mutation rate of the H strain of hepatitis C virus

NORIO OGATA*, HARVEY J. ALTER†, ROGER H. MILLER*, AND ROBERT H. PURCELL*

*Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, and †Department of Transfusion Medicine, Warren G. Magnuson Clinical Center, National Institutes of Health, Bethesda, MD 20892

Contributed by Robert H. Purcell, January 3, 1991

ABSTRACT **Patient H is an American patient who was infected with hepatitis C virus (HCV) in 1977. The patient became chronically infected and has remained so for the past 13 years. In this study, we compared the nucleotide and predicted amino acid sequences of the HCV genome obtained from plasma collected in 1977 with that collected in 1990. We find that the two HCV isolates differ at 123 of the 4923 (2.50%) nucleotides sequenced. We estimate that the mutation rate of the H strain of HCV is ≈1.92×10^{-3} base substitutions per genome site per year. The nucleotide changes were exclusively base substitutions and were unevenly distributed throughout the genome. A relatively high rate of change was observed in the region of the HCV genome that corresponds to the nonstructural protein 1 gene region of flaviviruses, where 44 of 960 (4.6%) nucleotides were different. Within this region there was a 39-nucleotide domain in which 28.2% of the nucleotides differed between the two isolates. In contrast, relatively few nucleotide substitutions were observed in the 5′ noncoding region, where only 2 of 276 (0.7%) nucleotides were different. Our results suggest that the mutation rate of the HCV genome is similar to that of other RNA viruses and that genes appear to be evolving at different rates within the virus genome.**

The principal etiologic agent of parenterally transmitted non-A, non-B (NANB) hepatitis recently was molecularly cloned from the plasma of an experimentally infected chimpanzee (1). The agent has been named hepatitis C virus (HCV). Tests for antibody to HCV (2) or, more directly, the HCV genome (3) clearly demonstrate that HCV is the major etiological agent of posttransfusion NANB hepatitis.

The HCV genome is a linear, positive-stranded RNA molecule of ≈10,000 nucleotides and encodes a polyprotein of ≈3000 amino acids (for a review, see ref. 4). Several stretches of amino acids in the HCV polyprotein share significant similarity with flavivirus- and pestivirus-encoded sequences (4–6). Therefore, HCV is considered to be distantly related to these virus groups.

It is known that the mutation rate of RNA viruses ranges from 10^{-1} to 10^{-4} base substitutions per genome site per year (7, 8). Thus, the amino acid sequence of virus proteins is subject to considerable change within a relatively short evolutionary period. This fact is generally regarded as an important feature in viral pathogenesis as well as in the establishment of diagnostic tools and the development of effective vaccines.

Recent studies indicate that the HCV genome is heterogeneous both among infected individuals (9–17) and within the same individual (9, 10, 13). In the present study, we compared the nucleotide sequence (and the predicted amino acid protein sequence) of the HCV genome obtained from plasma samples of a chronically infected patient over an interval of 13 years.‡ The results suggest that the HCV genome evolves rapidly and that different genes may evolve at different rates within the virus genome.

MATERIALS AND METHODS

Patient Studied. Patient H is an American with chronic posttransfusion NANB hepatitis (18). The plasma samples under study were obtained on July 12, 1977 (i.e., at the early acute stage of infection) and on August 1, 1990, which represents an interval of 13 years of continuous replication of the virus (P. Farci, H.J.A., and R.H.P., unpublished results). Virus from plasmapheresis performed in 1977 and 1990 was designated HCV isolate H77 and HCV isolate H90, respectively. There were no known episodes of reexposure to any blood or blood product during the 13-year interval.

RNA Preparation. RNA was extracted by a modification of the hot phenol method (19) as follows. Ten microliters of plasma was mixed with a solution containing 50 mM Tris·HCl (pH 8.0), 4.2 M guanidinium isothiocyanate, 0.5% (wt/vol) Sarkosyl, and 0.1 M 2-mercaptoethanol followed by an extraction with 1:1 (vol/vol) phenol/chloroform at 65°C. EDTA and $NaDodSO_4$ were added to a final concentration of 10 mM and 0.2% (wt/vol), respectively, to the aqueous phase, and the mixture was extracted with phenol/chloroform at room temperature. After precipitation with isopropanol, the RNA pellet was dissolved in 100 μl of water containing 10 mM dithiothreitol and 200 units of human placental ribonuclease inhibitor (rRNasin, Promega).

cDNA Synthesis. A 10-μl aliquot of the RNA solution was used for cDNA synthesis in a 20-μl reaction mixture containing 50 mM Tris·HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 1 mM of each of the four dNTPs, and 100 ng of random hexadeoxynucleotides. The mixture was heated to 70°C for 5 min, chilled on ice, and then treated with 200 units of Moloney murine leukemia virus reverse transcriptase (Superscript, BRL) at 42°C for 60 min.

Polymerase Chain Reaction (PCR). PCR was performed with "nested" (external and internal) primer pairs (20). A 5-μl aliquot of the cDNA reaction mixture was used for the first round of PCR amplification in a reaction volume of 100 μl containing 10 mM Tris·HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 400 μM of each of the four dNTPs, 0.5 μM of the external sense and antisense oligodeoxynucleotide primers, and 2.5 units of *Taq* DNA polymerase (Amplitaq, Perkin–Elmer/Cetus). The amplification reaction was performed by using an automatic thermocycler (Perkin–Elmer/Cetus) for 25–40 cycles of denaturation at 94°C for 1 min, primer annealing at 55°C for 2 min, and primer extension at 72°C for 3 min. For the second round of PCR amplification, 2–5 μl of the first PCR product was amplified exactly as described above except for the substitution of internal sense

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HCV, hepatitis C virus; NANB hepatitis, non-A, non-B hepatitis; PCR, polymerase chain reaction; NS1–NS5, nonstructural proteins 1–5.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M62381–M62386).

and antisense primers. Oligonucleotide primers were designed according to published HCV sequences (11, 16, 17) and were synthesized by using a DNA synthesizer (Applied Biosystems model 391). The location of the amplified fragments on the HCV genome is shown in Fig. 1. The size of each doubly amplified DNA fragment was confirmed by 2% agarose gel electrophoresis followed by ethidium bromide staining.

Direct DNA Sequencing. Amplified DNA was fractionated by agarose gel electrophoresis and purified by electroelution, Sephadex G-50 chromatography, and ethanol precipitation. Approximately 100 ng of DNA was used for each set of sequencing reactions by the dideoxynucleotide chain-termination method (21) with phage T7 DNA polymerase (Sequenase, United States Biochemical). Sequencing primers were either one of the internal primers used for PCR amplification or one of 32 primers synthesized to be complementary to the nucleotide sequence of the H strain of HCV. DNA sequencing was performed on at least two samples that were independently reverse-transcribed and PCR-amplified. It should be noted that of the 9846 bases sequenced, 8 nucleotides (6 in isolate H77 and 2 in isolate H90) repeatedly showed the presence of two bases at given positions: one major and a second minor base. This suggests that several regions of the H77 and H90 genomes exhibit sequence heterogeneity. Thus, the reported sequence is the consensus sequence of two to seven sequencing reactions.

RESULTS

Plasma was collected over a 13-year period from patient H who developed chronic, posttransfusion NANB hepatitis in 1977 (18). Recent investigation indicates that the etiological agent of the NANB hepatitis infection was HCV (P. Farci, H.J.A., and R.H.P., unpublished data). In the present study, 11 segments of the HCV genome were converted to cDNA, PCR-amplified, and used for nucleotide sequence analysis (Fig. 1). According to nucleotide sequence alignments of a published HCV sequence (17) with flavivirus and pestivirus sequences and comparisons of their predicted amino acid sequences, we tentatively define the 5′ noncoding region as positions −276 to −1; the core, or nucleocapsid, protein gene as positions 1–420; the envelope protein gene as positions 421–870; the NS1 gene (*NS1*) as positions 871–1830; and a part of the *NS2* gene as positions 1831–2334. When compared to pestivirus gene sequences, these domains correspond respectively to the 5′ noncoding region and regions encoding p20, gp62, gp53, and a 30-kDa product. Segments A–F in Fig. 1 cover 2610 bases from nucleotide position −276 through position 2334. Segments K to M in Fig. 1 consist of 1284 nucleotides from position 3625 to position 4908 and correspond to regions encoding a subregion of NS3 of flaviviruses or p80 (p125) of pestiviruses encompassing the helicase-like domain. Segments V and W in Fig. 1 consist of 1029 nucleotides from position 7579 to position 8607 and correspond to regions encoding a portion or NS5 of flaviviruses or p75 (p133) of pestiviruses encompassing the amino acids at the putative active site of the viral replicase. Thus, 4923 bases, or approximately half of the predicted genome length, of HCV were compared and are presented in Fig. 2.

Analysis indicates that the two HCV isolates show diversity at 123 of 4923 nucleotide positions (2.50%). Thus, the mutation rate of the H strain of HCV is estimated to be $\approx 1.92 \times 10^{-3}$ base substitutions per genome site per year. The nucleotide changes were exclusively base substitutions and were not uniformly distributed throughout the genome (Table 1). A relatively high rate of change was observed in *NS1* where 44 of 960 (4.6%) nucleotides were different. In contrast, relatively few nucleotide substitutions were observed in the 5′ noncoding region and the core gene, where 2 of 276 (0.7%) and 6 of 420 (1.4%) of the nucleotides, respectively, were different. Within *NS1*, a 39-nucleotide region (i.e., positions 1183–1221) had an extremely high mutation rate, where 11 (28.2%) nucleotide positions were associated with differences.

Overall, the ratio of transitions (i.e., pyrimidine-to-pyrimidine or purine-to-purine changes) to transversions (i.e., pyrimidine-to-purine or purine-to-pyrimidine changes) is 3.92. Interestingly the 39-nucleotide segment in *NS1* shows an extremely low ratio of 1.10, with only six transitions versus five transversions occurring in this region. A second region with a relatively low ratio is in the 5′ one-third of the sequenced *NS5* gene. The rates of base exchange in the coding regions of the HCV genome are 25.6% in the first codon position, 12.4% in the second, and 62.0% in the third position. In the 39-nucleotide region in *NS1*, base substitutions occur only in the first and the second codon positions.

Analysis of the predicted amino acid sequence corresponding to the two HCV isolates indicates that 40 of 1549 (2.58%) of the residues are different. As is clear from Table 1, these changes are not randomly distributed. NS1 has the most changes, where 23 of 320 (7.2%) residues are different. This high rate of change is due mainly to changes within the 39-nucleotide region of *NS1* described earlier, where 9 of 13 (69.2%) codons encode different (i.e., changed) amino acids. Other gene regions encode significantly lower amino acid differences that range from 1.2% in NS3 to 1.8% in NS2. The regions of NS3 and NS5 that have been shown previously to share sequence similarity with other virus groups are well conserved in the two isolates of HCV (H77 and H90) examined here (unpublished data). Overall, 22 of 40 (55.0%) codon changes between H77 and H90 result in nonconservative amino acid changes. This is especially true in the *NS1* gene region, where the rate of the nonconservative change is 65.2% (i.e., 15 of 23). Thus, the genome sequence of HCV from a chronically infected patient has undergone significant change during 13 years of virus replication.

Comparison of the predicted amino acid sequences of the putative core, envelope, and a part of NS1 encoded by H77 and H90 with those of other HCV isolates (Fig. 3) demonstrates that H77 and H90 are more closely related to each other than they are to any other isolate. This relationship is prominent in the carboxyl terminus of the envelope protein through the amino terminus of NS1 when compared with predicted amino acid sequences of other isolates. The variable region of NS1, corresponding to amino acids 395–407 (Fig. 3), is also extremely variable in the predicted products of other HCV isolates. Thus, a defined region within the predicted *NS1* gene protein of HCV exhibits hypervariability among all isolates examined.

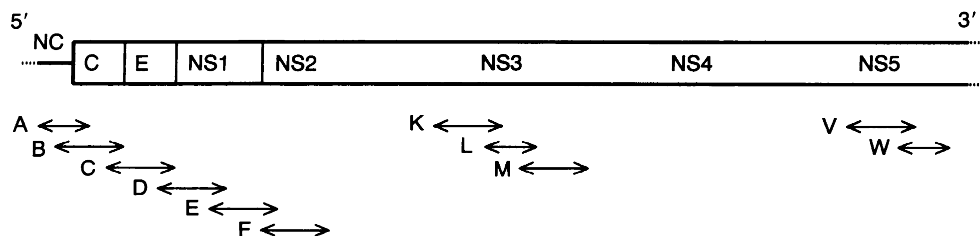

FIG. 1. Regions of the HCV genome amplified by PCR. The map of the HCV genome illustrating the noncoding (NC) region and regions encoding the core (C), envelope (E), and nonstructural proteins 1–5 (NS1–NS5) is presented. Arrows define the limits of the genome regions amplified by PCR and sequenced.

DISCUSSION

Recent data from other laboratories suggest that the HCV genome sequence is heterogeneous within a given isolate (9, 10, 13). In the light of this observation, a desirable way for the analysis of viral genomes among isolates is to obtain the "average" or "consensus" sequence in each isolate. Although direct sequencing of the viral RNA genome is the ideal method of analysis, the approach is not practical. Therefore, we used direct sequencing of reverse-transcribed, PCR-amplified HCV genomic RNA. Several independently obtained samples were analyzed for each genome region to circumvent any potential artifacts that might arise from misincorporation during the amplification reactions (22, 23).

We define the mutation rate of the HCV genome as the number of nucleotide substitutions that occurred per genome site per year of continuous replication of the virus in patient H. It is important to note that this definition does not take into account the replication cycle time of HCV, which is unknown at the present time and is dependent upon the ability of mutant progeny virus to be perpetuated in the virus population (8). Using this definition, we estimate the mutation rate of the HCV genome to be $\approx 10^{-3}$ nucleotide substitutions per site per year. This value is similar to that of other RNA viruses that use a polymerase that lacks a proofreading function for replication of the virus genome (24–26) and is about a million-fold higher than the rate observed for the replication of prokaryotic and eukaryotic chromosomal DNA (7, 8).

Comparing the nucleotide sequence of the H77 and H90 isolates to other published HCV isolates confirms that the greatest conservation of sequence occurs in the 5′ noncoding region. Although the function of this region is not known at present, we find that two nucleotide domains share statistically significant sequence similarity with pestivirus 5′ noncoding region sequences (R.H.M., unpublished observation; ref. 33). These findings suggest that this region may have some structural or functional significance in virus replication or gene expression.

Comparison of the predicted amino acid sequences of the H77 and H90 isolates to those of other HCV isolates indicates that the greatest similarity occurs within the central region and carboxyl terminus of NS3 as well as the carboxyl terminus of NS5. The greatest degree of conservation occurs within the NS3 and NS5 domains shown by us to share sequence similarity with predicted products of flaviviruses, pestiviruses, and carmoviruses (5). The conserved region of NS3 possesses sequence similarity with the active site of helicase-like nucleoside triphosphate-binding proteins, while the conserved region of NS5 may contain the amino acids composing the catalytic site of the virus replicase (5). It is likely that these sequences are conserved to keep the functional domains of these proteins intact.

It is noteworthy that there is perfect conservation of 30 cysteine residues in the structural gene products and in NS1 and NS2. This finding implies that there is little, if any, change in the three-dimensional structure of these proteins.

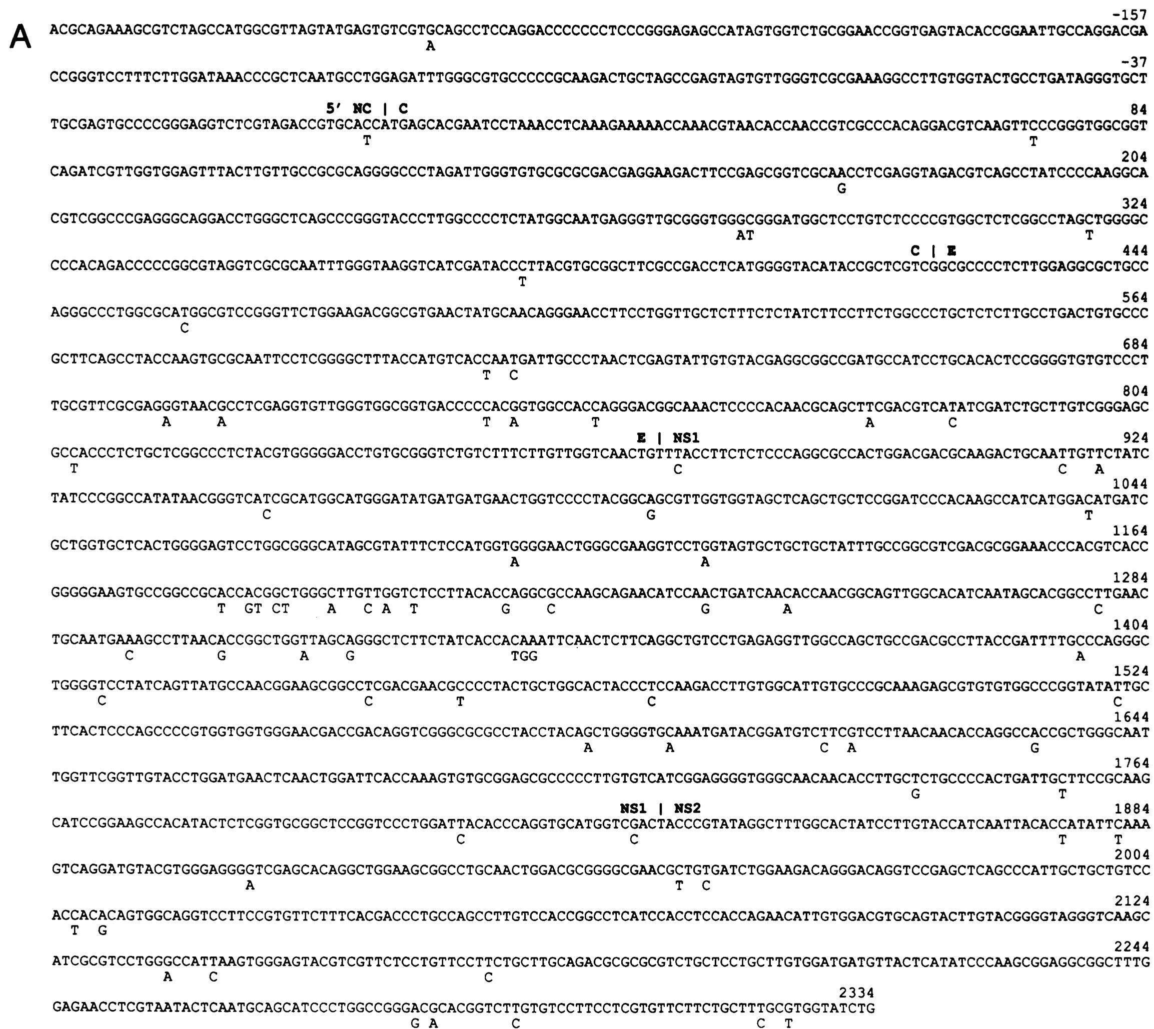

FIG. 2. (*Figure continues on the opposite page.*)

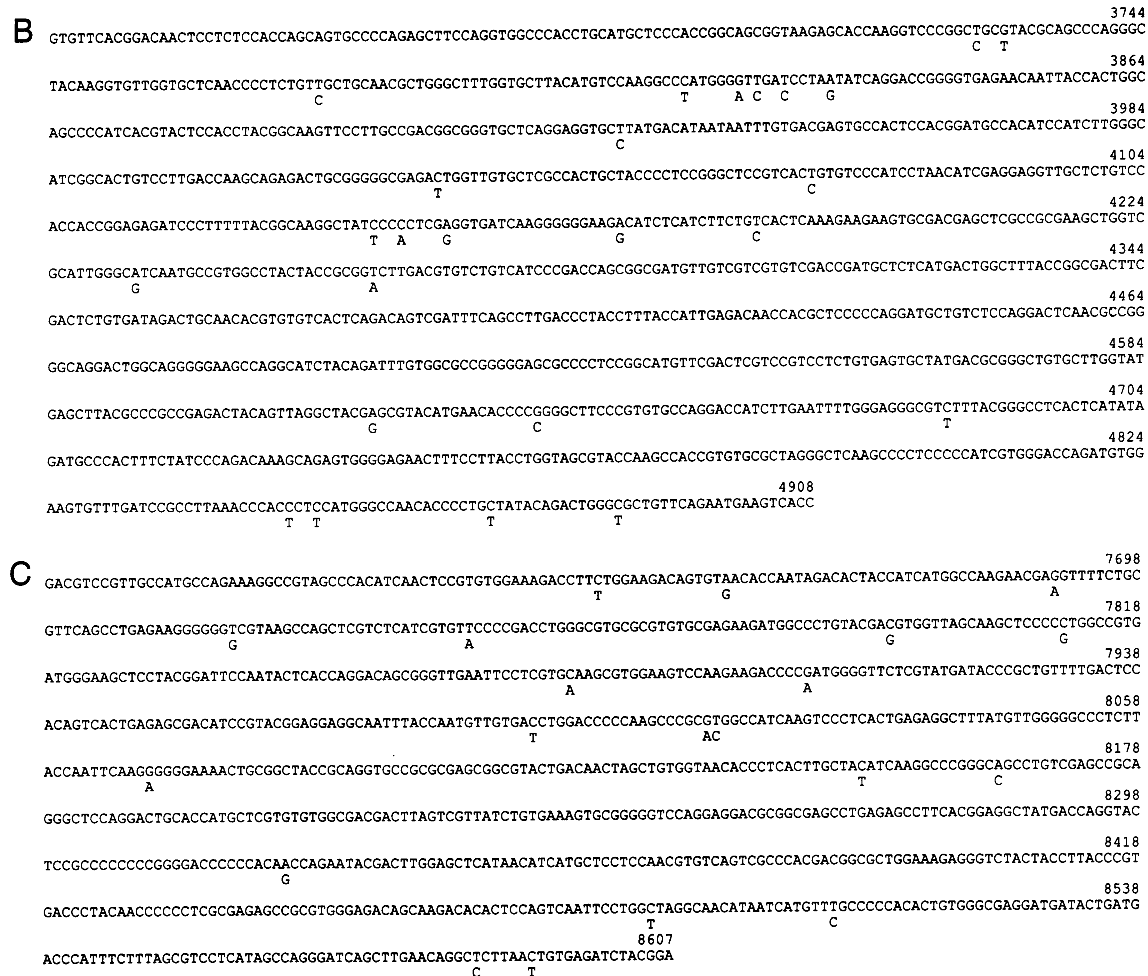

FIG. 2. Nucleotide sequence of the H strain of HCV. In each row, the upper line represents the sequence of isolate H77, while the lower line depicts the changes found in isolate H90 in the 5′ noncoding region (5′ NC) and the core (C), envelope (E), *NS1*, and *NS2* gene regions (*A*); the *NS3* region (*B*); and the *NS5* region (*C*). Map unit designations are taken from a previously published HCV sequence (17). The 5′ noncoding region (5′ NC) and the approximate boundaries of the core, envelope, *NS1*, and *NS2* gene regions are shown.

In addition, the potential N-glycosylation sites of these proteins are also well preserved. Thus, it is likely that functional and/or structural constraints have played a role in minimizing change in these regions during the course of virus evolution.

The extremely high rate of nucleotide changes in *NS1* (and predicted amino acid changes in its product), especially within a 39-nucleotide region, can be contrasted with the above findings. Although the exact boundary of this gene and its function in virus replication are not known, the abundance of possible N-glycosylation sites in its product suggests that this segment encodes a glycoprotein. The high proportion of codon changes in this domain suggests that the functional/structural constraints are very low. The analogous gene of flaviviruses codes for a nonstructural protein that is expressed on the surface of infected cells (27, 28). However, despite its nonstructural nature, it is important in stimulating resistance to illness, perhaps through a cell-mediated immune mechanism. In pestiviruses the analogous gene is thought to code for an additional envelope protein (29, 30). In both cases, the proteins are exposed to the host's immune response and its attendant selective pressure in favor of antigenically distinct variants. The high percent of nucleotide substitutions resulting in amino acid changes supports this conclusion. Thus, the hypervariable region of the NS1 protein of HCV may share similarities with the hypervariable region of the envelope protein of human immunodeficiency virus (HIV), an immunodominant domain involved in strain-specific neutralization of the virus (25). Although the exact nature of the *NS1* gene product of HCV is not yet known, it is very likely that its hypervariable region is the product of selective immunological pressures.

Table 1. Comparison of nucleotide changes and predicted amino acid differences

Region	Changes (%) Nucleotide	Amino acid	% nt → AA*
5′ noncoding	2/276 (0.7)	NA	NA
Core	6/420 (1.4)	2/140 (1.4)	33.3
Envelope	11/450 (2.4)	2/150 (1.3)	18.2
NS1	44/960 (4.6)	23/320 (7.2)	52.3
NS2	15/504 (3.0)	3/168 (1.8)	20.0
NS3	25/1284 (1.9)	5/428 (1.2)	20.0
NS5	20/1029 (1.9)	5/343 (1.5)	25.0
Total	123/4923 (2.5)	40/1549 (2.6)	32.5

NA, Not applicable.
*Percent of nucleotide (nt) changes that result in amino acid (AA) changes.

The genetic variation between H77 and H90 isolates was presumably generated during RNA replication and accentuated by the lack of a proofreading activity in the RNA-dependent RNA polymerase (7, 8). We found that the only differences between the two isolates were base substitutions. In studies by others, a bias toward base transitions, possibly dictated by G·U mispairing, has been shown to occur during

FIG. 3. Alignment of the amino acid sequences in single-letter code predicted from the core, envelope, and *NS1* gene regions of five isolates of HCV. The predicted amino acids of the core (positions 1–140), envelope (141–290), and part of NS1 (291–440) are aligned for the following isolates: 1, H77; 2, H90; 3, isolate from ref. 14; 4, HC-J4 (11); 5, HC-J1 (11); and 6, isolate from ref. 17. It should be noted that gene regions are approximate as discussed in the text. The hypervariable 39-nucleotide region discussed in the text corresponds to amino acid positions 395–407. The approximate boundaries of the core (C), envelope (E), and NS1 proteins are shown.

replication by RNA polymerases (31, 32). In this respect, a relatively low proportion of transition to transversion in *NS1*, particularly in the 39-nucleotide segment, is intriguing. A second interesting feature of the base changes in *NS1*, especially in the 39-nucleotide segment, is the substantial number of changes at the first and the second codon positions. Thus, it seems possible that not only selective forces but also a different frequency or mode of replicative errors in the growing RNA chains may have influenced the "segmental" evolution of HCV.

In summary, we have presented evidence for the rapid evolution of the HCV genome in a chronically infected patient followed for 13 years. We have also demonstrated that regions of the HCV genome appear to be evolving at different rates. These findings could, in part, explain the large diversity of HCV genomes among isolates from different individuals.

1. Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. (1989) *Science* **244**, 359–362.
2. 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., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W.-S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W. & Houghton, M. (1989) *Science* **244**, 362–364.
3. Weiner, A. J., Kuo, G., Bradley, D. W., Bonino, F., Saracco, G., Lee, C., Rosenblatt, J., Choo, Q.-L. & Houghton, M. (1990) *Lancet* **335**, 1–3.
4. Choo, Q.-L., Weiner, A. J., Overby, L. R., Kuo, G., Houghton, M. & Bradley, D. W. (1990) *Br. Med. Bull.* **46**, 423–441.
5. Miller, R. H. & Purcell, R. H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2057–2061.
6. Collett, M. C., Anderson, D. K. & Retzel, E. (1988) *J. Gen. Virol.* **69**, 2637–2643.
7. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandelPol, S. (1982) *Science* **215**, 1577–1585.
8. Domingo, E. & Holland, J. J. (1988) in *RNA Genetics*, eds. Domingo, E., Holland, J. J. & Ahlquist, P. (CRC, Boca Raton, FL), Vol. 3, pp. 3–36.
9. 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) *Nucleic Acids Res.* **24**, 10367–10372.
10. Kato, N., Ohkoshi, S. & Shimotohno, K. (1989) *Proc. Jpn. Acad.* **65**, 219–223.
11. Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T., Yoshizawa, H., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1990) *Jpn. J. Exp. Med.* **60**, 167–177.
12. Maeno, M., Kaminaka, K., Sugimoto, H., Esumi, M., Hayashi, N., Komatsu, K., Abe, K., Sekiguchi, S., Yano, M., Mizuno, K. & Shikata, T. (1990) *Nucleic Acids Res.* **18**, 2685–2689.
13. Takeuchi, K., Boonmar, S., Kubo, Y., Katayama, T., Harada, H., Ohbayashi, A., Choo, Q.-L., Kuo, G., Houghton, M., Saito, I. & Miyamura, T. (1990) *Gene* **91**, 287–291.
14. Takeuchi, K., Kubo, Y., Boonmar, S., Watanabe, Y., Katayama, T., Choo, Q.-L., Kuo, G., Houghton, M., Saito, I. & Miyamura, T. (1990) *Nucleic Acids Res.* **18**, 4626.
15. Enomoto, N., Takada, A., Nakao, T. & Date, T. (1990) *Biochem. Biophys. Res. Commun.* **170**, 1021–1025.
16. Houghton, M., Choo, Q.-L. & Kuo, G. (1988) Eur. Patent Appl. 88,310,922.5 and Publ. 318,216.
17. Houghton, M., Choo, Q.-L. & Kuo, G. (1990) Eur. Patent Appl. 90,302,866 and Publ. 388,232.
18. Feinstone, S. M., Alter, H. J., Dienes, H. P., Shimizu, Y., Popper, H., Blackmore, D., Sly, D., London, W. T. & Purcell, R. H. (1981) *J. Infect. Dis.* **144**, 588–598.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
20. Kaneko, S., Unoura, M., Kobayashi, K., Kuno, K., Murakami, S. & Hattori, N. (1990) *Lancet* **335**, 976.
21. Bachmann, B., Luke, W. & Hunsmann, G. (1990) *Nucleic Acids Res.* **18**, 1309.
22. Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A. & Kunkel, T. A. (1989) *Mol. Cell. Biol.* **9**, 469–476.
23. Eckert, K. A. & Kunkel, T. A. (1990) *Nucleic Acids Res.* **18**, 3739–3744.
24. Buonagurio, D. A., Nakada, S., Parvin, J. D., Krystal, M., Palese, P. & Fitch, W. M. (1986) *Science* **232**, 980–982.
25. Hahn, B. H., Shaw, G. M., Taylor, M. E., Redfield, R. R., Markham, P. D., Salahuddin, S. Z., Wong-Staal, F., Gallo, R. C., Parks, E. S. & Parks, W. P. (1986) *Science* **232**, 1548–1553.
26. Imazeki, F., Omata, M. & Ohto, M. (1990) *J. Virol.* **64**, 5594–5599.
27. Gibson, C. A., Schlesinger, J. J. & Barrett, A. D. T. (1988) *Vaccine* **6**, 7–9.
28. Yasuda, A., Kimura-Kuroda, J., Ogimoto, M., Miyamoto, M., Sata, T., Sato, T., Takamura, C., Kurata, T., Kojima, A. & Yasui, K. (1990) *J. Virol.* **64**, 2788–2795.
29. Stark, R., Rumenapf, T., Meyers, G. & Thiel, H.-J. (1990) *Virology* **174**, 286–289.
30. Weiland, E., Stark, R., Haas, B., Rumunapf, T., Meyers, G. & Thiel, H.-J. (1990) *J. Virol.* **64**, 3563–3569.
31. Jou, W. M. & Fiers, W. (1976) *J. Mol. Biol.* **106**, 1047–1060.
32. Both, G. W., Sleigh, M. J., Cox, N. J. & Kendal, A. P. (1983) *J. Virol.* **48**, 52–60.
33. Han, J. H., Shyamala, V., Richman, K. H., Brauer, M. J., Irvine, B., Urdea, M. S., Tekamp-Olson, P., Kuo, G., Choo, Q.-L. & Houghton, M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1711–1715.