Vol. 189, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 30, 1992 Pages 119-127

CHARACTERIZATION OF HYPERVARIABLE REGIONS IN THE PUTATIVE ENVELOPE PROTEIN OF HEPATITIS C VIRUS

Nobuyuki Kato [1], Yuko Ootsuyama [1], Showgo Ohkoshi [2], Takahide Nakazawa [1], Hitomi Sekiya [1], Makoto Hijikata [1], and Kunitada Shimotohno [1]

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

[2] Third Department of Internal Medicine, School of Medicine, Niigata University, 1-757 Asahimachi, Niigata-city 951, Japan

Received September 29, 1992

SUMMARY: We previously identified two hypervariable regions[HVR1(27 amino acids) and HVR2(7 amino acids)] in the putative envelope glycoprotein(gp70) by comparison of the amino acid sequences of many isolates of the HCV-II genotype. To understand the functional features of these HVRs, using the polymerase chain reaction we analyzed the rate of actual sequence variability in the region including HVR1 and HVR2 of HCV isolated successively at intervals of several months from two patients with chronic C-type hepatitis. In both patients, the amino acid sequence of HVR1, but not HVR2, was found to change dramatically during the observation period(about one amino acid per month). However, no alteration of the amino acid sequence of HVR1 of HCV was observed in a patient in the acute phase of chronic hepatitis. Restriction digestion analysis of sequence diversity showed that a HCV genome with a newly introduced mutation in HVR1 often became the predominant population at the next time of examination. Alterations of amino acids in HVR1 occurred sequentially in the two patients in the chronic phase. These findings suggest that mutations in HVR1 are involved in the mechanism of persistent chronic HCV infection.

Hepatitis C virus(HCV) is responsible for most cases of posttransfusional non-A, non-B(NANB) hepatitis throughout the world (1,2). Molecular cloning of the HCV genome revealed that there are at least six different genotypes(HCV-I to HCV-VI) in the world

Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under Accession Nos. D12942 and D12972.

0006-291X/92 $4.00

(3). The entire HCV genomes of HCV-I to HCV-IV have been molecularly cloned(4-7). The HCV genome is a positive-stranded RNA molecule of about 9.5 kilobases and encodes a large polyprotein precursor of about 3000 amino acids. The amino acid sequence of the polyprotein revealed that HCV belongs to a new genus distinct from the flaviviruses and pestiviruses of the *Flaviviridae* (4).

A characteristic feature of the HCV genome is the sequence diversity among virus isolates. About 20 to 35% of the nucleotide sequence of the virus genome differs in different genotypes. Furthermore, different isolates of the viral genome belonging to the same genotype showed 5 to 8% differences in nucleotide sequences and 4 to 5% differences in amino acid sequences(4,8). These sequence diversities are distributed along the whole genome except for the conserved 5' noncoding region, but the putative envelope region encoding gp35 and gp70 shows especially marked sequence diversity(8,9). By comparing many HCV isolates (HCV-I and -II genotypes), we identified two hypervariable regions[HVR1(27 amino acids) and HVR2(7 amino acids)] in the N-terminal region of gp70 of the HCV-II genotype(9-11) which is the major genotype in Japan(12). HVR1(21 amino acids) in the same region as in the HCV-II genotype is also present in the HCV-I genotype, but HVR2 seldom shows sequence variation in the HCV-I genotype(11,13). Two groups have observed sequence variability of 8 and 9 amino acids in HVR1 of the HCV-I genotype during periods of 13 and 2 years by comparing amino acid sequences(14,15). However, their results may not reflect the actual rate of variability in HVR1, because their comparisons were made at only two time points separated by long intervals. To obtain more precise information on the rate of variability in HVRs, we therefore analyzed the HVRs of the HCV-II genotype in two patients with C-type hepatitis at intervals of a few months.

MATERIALS AND METHODS

Patients. Patient NY: A man of 28 years old with chronic NANB hepatitis without a history of blood transfusion. He visited the hospital of Niigata University in March, 1988 and received interferon therapy(total 84MU of interferon β) for two months from April, 1988. However, the disease was not cured and has remained in the chronic stage. Patient RS: A woman of 22 years old, diagnosed as having acute NANB hepatitis on her first visit to hospital. The onset of hepatitis was in March, 1991 without blood transfusion and

developed to the chronic stage. The sampling times of blood and the profile of alanine aminotransferase(ALT) are shown in Figure 1.

cDNA synthesis and PCR amplification. RNAs from plasma were prepared as described(4). RNA from 10 μl of plasma was used for cDNA synthesis and PCR amplification(RT-nested PCR) as described previously(11).

cDNA cloning and sequencing. PCR products were cloned into the *Eco* RI site of the pTZ19R plasmid vector as described previously(11). Nucleotide sequences were determined by the dideoxy-nucleotide chain-termination method using a DNA sequencing kit(USB Corp., OH.) or A.L.F. DNA sequencer(Pharmacia).

Restriction enzyme analysis of PCR products. A sample of 10 μl of PCR products from patient NY was extracted with phenol:chloroform before ethanol precipitation. The DNA fragments obtained were digested with a restriction enzyme, *Bsa* JI, and the digests were separated in 3% agarose gel.

RESULTS AND DISCUSSION

Four plasma specimens from patient NY and five plasma specimens from patient RS were obtained at intervals of several months(see arrows in Fig.1). The RNAs from these specimens were used to detect the putative envelope region containing HVRs of the HCV genome. RT-nested PCR was carried out as described previously(11). A distinct band of the expected size[484 bp containing positions 1434 to 1847 of HCV-J(4)] was detected in all specimens from both patients by staining with ethidium bromide after agarose gel electrophoresis (Fig.2). The amplified putative envelope region encodes the C-terminal 15 amino acids of gp35 and the N-terminal 123 amino acids of gp70, which correspond to positions 369 to 506 of the HCV-II polyprotein precursor(10). The amplified 484 bp cDNA fragment from each specimen was isolated and digested with *Eco* RI and then inserted into the *Eco* RI site of pTZ19R plasmid vector. The nucleotide sequences of three to five independent cDNA clones from each specimen were determined to obtain precise information on microheterogeneities, as observed previously(11), and to determine the consensus sequence of each specimen. The amino acid sequences deduced from the nucleotide sequences of cDNAs are summarized in Figures 3 and 4. The results obtained from patient NY with typical chronic NANB hepatitis revealed that the amino acid sequence of HVR1 changed rapidly and constantly with time(Fig. 3). The consensus

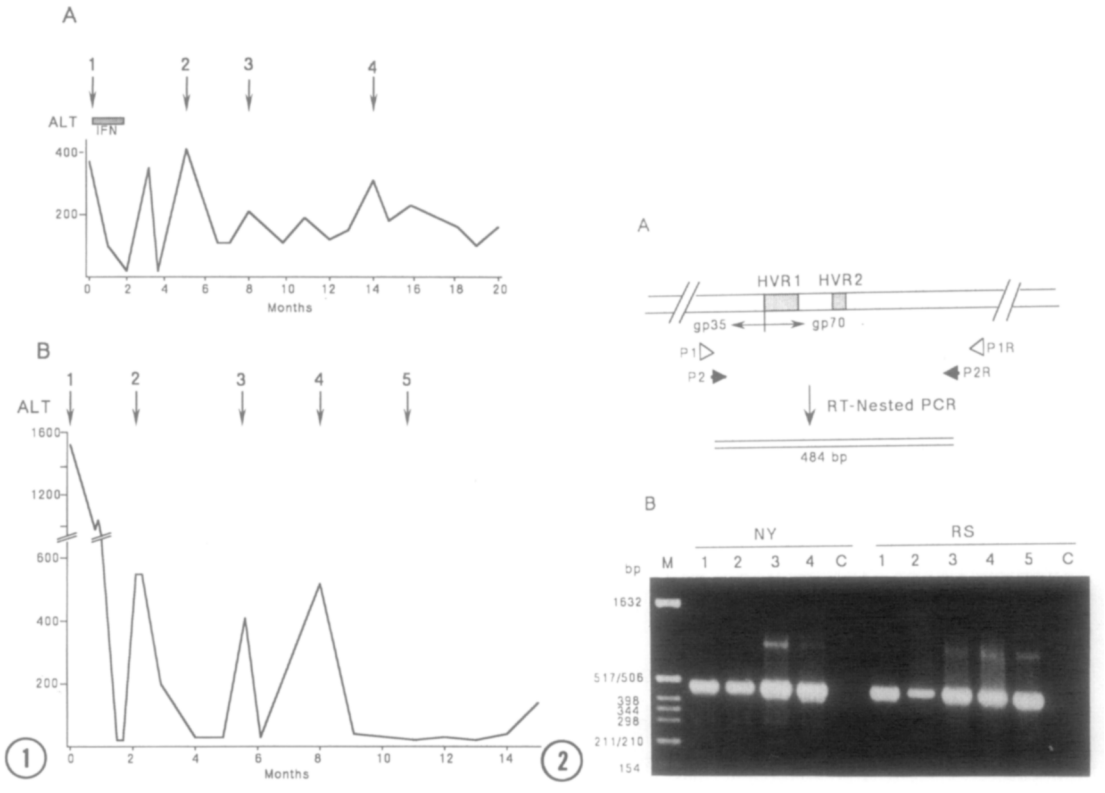

Figure 1. ALT profiles and sampling times of patients NY(A) and RS(B).The profiles of ALT values(units/liter) for patient NY from May, 1988 to January, 1990 and for patient RS from March, 1991 to June, 1992 are shown in (A) and (B), respectively. The arrows denote the times when plasma specimens were obtained for extraction of HCV RNA. The shaded bar indicates the period of interferon(IFN) treatment.

Figure 2. Region amplified by RT-nested PCR and detection of HCV RNA in plasma from patients NY and RS. The region amplified in this study is shown schematically in (A). RT-nested PCR products were detected by staining with ethidium bromide after 3% agarose gel electrophoresis (B). Lane M, *Hin* fI digests of pBR322 as size markers. Numbers of lanes 1 to 4 of NY and lanes 1 to 5 of RS correspond to sample numbers of patients in Figure 1.

sequences of the NY1 series and NY2 series differed in seven amino acids. These specimens were obtained at an interval of five months. Results on comparison of NY2 and NY3, and NY3 and NY4 were similar: in these cases, five and eight amino acids in HVR1 were changed in intervals of three and six months, respectively. In contrast to HVR1, the amino acid sequences of HVR2 were similar in NY1 to NY4. Results for patient RS in the chronic phase about 6 months after the onset of disease were similar to those in patient NY(Fig.4).

HVR1

```
              369                                                                      440
NY1 Cons.     AK VLIVMLLFAG VDGSTRVTGG QQGRAVHGIA SLFSLGASQK IQLVNTNGSW HINRTALNCN DSLQTGFLAA
  1-1         -- ---------- ---------- ---------- ---------- ---------- ---------- ----------
  1-2         -- ---------- ---------- ------Q-F- ---R-----E ---------- ---------- ----------
  1-3         -- ----V----- ---------- ---------- ---------- N--------- ---------- ----------
  1-4         V- ---------- ---------- ---------- ---------- ---------- ---------- ----------

NY2-1         -- ---------- ---------- ---H-A-SLT ---R-----N ---------- ---------- ----------
  2-2         -- A--------- ---------- ---H-A-SLT ---R-----N ---------- ---------- G---------
  2-3         -- ---------- ---------- ---H-A-SLT ---R-----N ---------- ---------- ----------

NY3-1         -- ---------- ---N------ R----A-SLT ----P----N ---------- ---------- ----------
  3-2         -- ---------- A--N------ R----A-SLT ----P----N ---------- ---------- ----------
  3-3         -- ---------- ---G------ R----A-SLT ----P---EN -R-------- ---------- ----------
  3-4         -- ---------- -------S-- -----A-SLT ---T-----N ---------- ---------- ----------
  3-5         -- ---------- -----H---A L----AY--T -FL-H-P--- ---------- ---------- ----------

NY4-1         -- ---------- -----Q-M-- -----AYSL- --L-P--N-- ---------- ---------- ----------
  4-2         -- ---------- -----Q-M-- -----AYSL- --LGP----- ---------- --------R- ----------
  4-3         -- ---------- -----Q-M-- -----AYSL- --L-P----- ---------- ---------- ----------
```

HVR2

```
                 441                                                              506
NY1 Cons.        LFYSNKFNAS GCPERMASCR PIDKFAQGSG AITYVVPNNL DQRPYCWHYA PQPCGIIPAS QVCGPV
  1-1            ---------- ---------- ---------- ---------- ---------- ------M--- ------
  1-2            ---------- ---------- ---------- ---------- G--------- -----V---- ------
  1-3            ---------- ---------- ---------- ---------- ---------- ---------- ------
  1-4            ---------- ---------C ---------- ---------- ---------- ---------- ------

NY2-1            ---------- ---------- ---------- ---------- G--------- ---------- ------
  2-2            -S-------- ---------- ---------- ---------- ---------- ---------- ------
  2-3            ---------- ---------- ---------- ---------- ---------- ---------- ------

NY3-1            ------Y--- ---------- ---------- ---------- ---------- ---------- ------
  3-2            ---------- ---------- ---------- ---------- ---------- ---------- ------
  3-3            ----T--D-- ---------- ---------- ---------- ---------- ---------- ------
  3-4            ---------- ---------- ---------- ---------- ---------- ---------- ------
  3-5            ---------- ---------- ---------- ---------- ---------- ---------- ------

NY4-1            ---------- ---------- ---------- ---------- ---------- ---------- --R---
  4-2            ---------- -------C-- ---------- ---------- ---------- ---------- --R---
  4-3            ---------- ---------- ---------- ---------- ---------- ---------- --R---
```

Figure 3. Deduced amino acid sequences in the amplified region from patient NY. Positions 369 to 506[numbered according to Kato et al.(4)] are shown. Amino acid sequences are indicated by the single letter code. Consensus sequences of NY1 were obtained from those of NY1-1 to NY1-4. Capital letters indicate different amino acids from those in the consensus sequence of NY1. HVR1 and HVR2 are shown by lines above the sequences.

Alterations of the amino acid sequences were clustered in HVR1, and the first alteration (Phe to Leu in amino acid position 399) was observed about 6 months after the onset of disease. The alteration of amino acid sequences was remarkable and sequential in accordance with development of the chronic phase(Figure 4).

The relationship of quantitative change of HCV with the change of the sequence of HVR1 with time was examined by amplifying cDNA products from NY1 and NY2 and by digesting them with a restriction enzyme, *Bsa* JI. A *Bsa* JI site(CCNNGG) is present in the HVR1 of 4

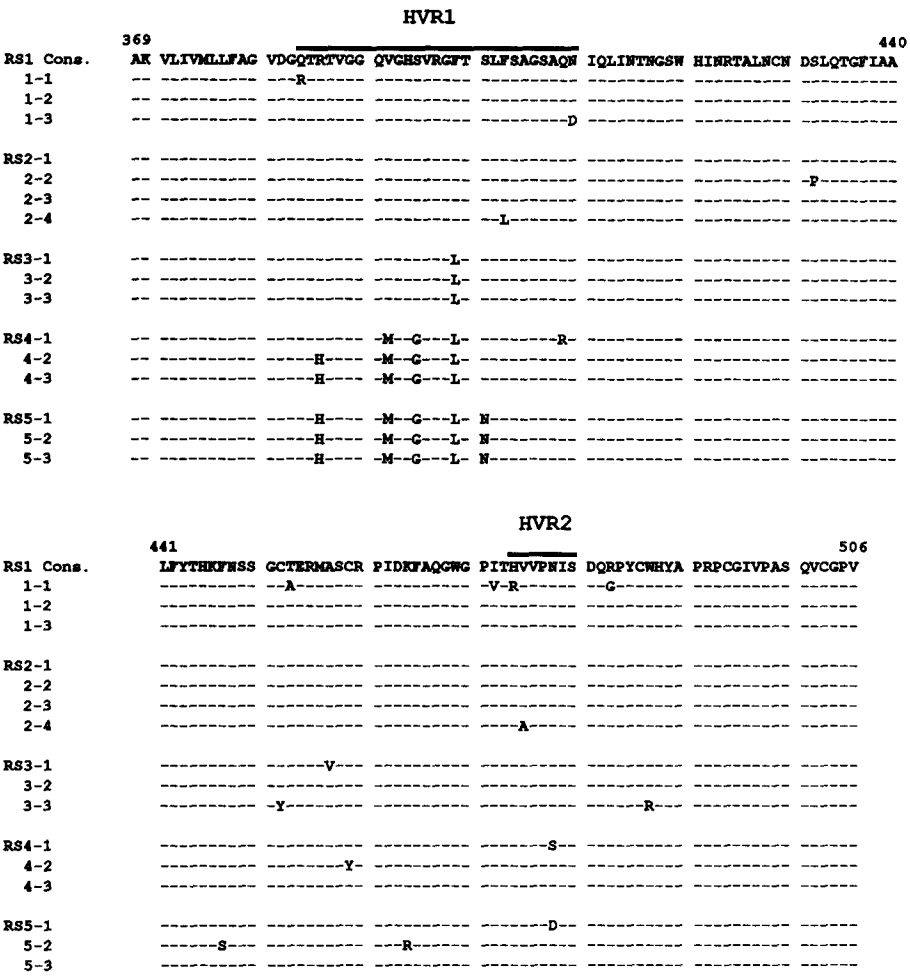

Figure 4. Deduced amino acid sequences in the amplified region from patient RS. Amino acid sequences from each cDNA clone are indicated as in Fig.3.

clones from NY1(CCACGG), but not in 3 clones from NY2(CCACAG). This substitution in the nucleotide sequence causes an alteration of Gly to Ser. If alternation in the nucleotide sequence in HVR1 is at a quantitative level in NY1 and NY2, the NY1 specimen should produce a 359 bp fragment and the NY2 specimen should produce a 441 bp fragment as the largest digestion product, as shown in Figure 5A. Figure 5B shows a typical result. We obtained a band of 359 bp from only the NY1 specimen and a band of 441 bp from only the NY2 specimen(Fig.5B) after digestion with *Bsa* JI. These results indicate that HCV with different sequences of HVR1 from previous dominant isolates became major species during the next few months.

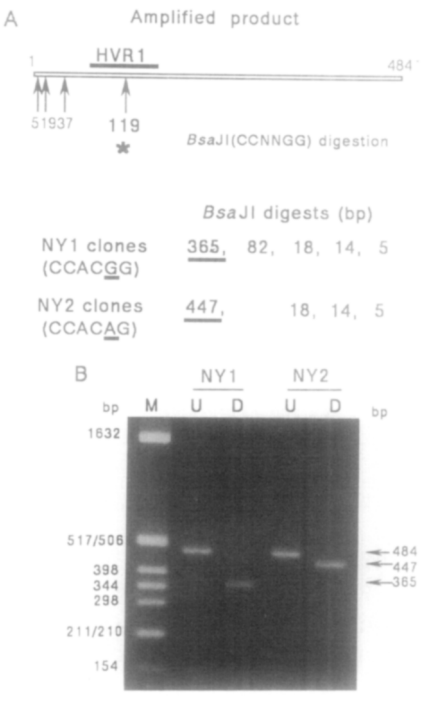

Figure 5. Analysis of sequence diversity in HVR1 from specimens NY1 and NY2. Fragments of the amplified product digested with *Bsa* JI are shown schematically in (A). *Bsa* JI digests of amplified products from NY1 and NY2 specimens are separated in 3% agarose gel and stained with ethidium bromide(B). Lane M, *Hin* fI digests of pBR322 as size markers. D, digest with *Bsa* JI; U, undigested sample.

From our results, the rate of alteration of amino acids in HVR1 is higher than expected from previous reports(14,15) on comparisons of only two time points of intervals of 13 and 2 years. The alteration of amino acids occurred sequentially in both NY and RS specimens strongly suggest that these alterations were due to direct mutations during viral replication, although the possibility that the changes were caused by quantitative change of HCVs with different sequences of HVRs is not completely ruled out.

The frequent variations of HVRs suggest that the observed amino acid sequences of HVR1 might result from constant selection by the immunosystem or might reflect the results of escape from immunosurveillance. Therefore, HVR1 may contain an epitope, which is specific for each virus isolate, against neutralizing antibodies to HCV. ALT values often vary in parallel with the level of viremia (16). This association may be explained by immunological surveillance of the virus by the host, which may have a cytopathic

effect on hepatocytes. Alternatively, immunological attack on virus infected cells by cytotoxic T-cell may result in elevation of ALT values. However, in this case increase in ALT values would not necessarily be associated with the condition of severe viremia. In this regards, it seems important to analyze the amount of virus in each serum sample examined in this study.

The present results suggest that the duration of C-type hepatitis may be related to mutations of HVR1 in HCV to escape from the immunosurveillance system. HVR1 must be examined as a possible epitope against virus specific antibodies.

ACKNOWLEDGMENTS: This study was supported by Grants-in-Aid for Cancer Research and for a Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare of Japan. T.N. and H.S. are recipients of Research Resident Fellowships from the Foundation for Promotion of Cancer Research.

REFERENCES

1. Choo,Q.-L., Kuo,G., Weiner,A.J., Overby,L.R., Bradley,D.W. and 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,W.-S., Lee,W. -S., Kuo,C., Berger,K., Shuster,J.R., Overby,L.R., Bradley,D.W. and Houghton,M. (1989) Science 244, 362-364.
3. Mori,S., Kato,N., Yagyu,A., Tanaka,T., Ikeda,Y., Petchclai,B., Chiewsilp,P., Kurimura,T. and Shimotohno,K. (1992) Biochem. Biophys. Res. Commun. 183, 334-342.
4. Kato,N., Hijikata,M., Ootsuyama,Y., Nakagawa,M., Ohkoshi,S., Sugimura,T. and Shimotohno,K. (1990) Proc.Natl. Acad. Sci. USA 87, 9524-9528.
5. Choo,Q.-L., Richman,K.H., Han,J.H., Berger,K., Lee,C., Dong,C., Gallegos,C., Coit,D., Medina-Selby,A., Barr,P.J., Weiner,A.J., Bradley,D.W., Kuo,G. and Houghton,M. (1991) Proc. Natl. Acad. Sci. USA 88, 2451-2455.
6. Okamoto,H., Okada,S., Sugiyama,Y., Kurai,K., Iizuka,H., Machida,A., Miyakawa,Y. and Mayumi,M. (1991) J. Gen. Virol. 72, 2697-2704.
7. Okamoto,H., Kurai,K., Okada,S., Yamamoto,K., Lizuka,H., Tanaka,T., Fukuda,S., Tsuda,F. and Mishiro,S. (1992) Virology 188, 331-341.
8. Tanaka,T., Kato,N., Nakazawa,M., Ootsuyama,Y., Cho,M-J., Nakazawa,T., Hijikata,M., Ishimura,Y. and Shimotohno,K.(1992) Virus Res. 23, 39-53.
9. Hijikata,M., Kato,N., Ootsuyama,Y., Nakagawa,M., Ohkoshi,S. and Shimotohno,K.(1991) Biochem. Biophys. Res. Commun. 175, 220-228.
10. Hijikata,M., Kato,N., Ootsuyama,Y., Nakagawa,M. and Shimotohno,K. (1991) Proc. Natl. Acad. Sci.USA 88, 5547-5551.

11. Kato,N., Ootsuyama,Y., Tanaka,T., Nakagawa,M., Nakazawa,T., Muraiso,K., Ohkoshi,S., Hijikata,M. and Shimotohno,K. (1992) Virus Res. 22, 107-123.
12. Kato,N., Ootsuyama,Y., Ohkoshi,S., Nakazawa,T., Mori,S., Hijikata,M. and Shimotohno,K.(1991) Biochem. Biophys. Res. Commun. 181, 279-285.
13. Weiner,A.J., Brauer,M.J., Rosenblatt,J., Richman,K.H., Tung,J., Crawford,K., Bonino,F., Saracco,G., Choo,Q-L., Houghton,M. and Han,J.H. (1991) Virology 180, 842-848.
14. Ogata,N., Alter,H.J., Miller,R.H. and Purcell,R.H. (1991) Proc. Natl. Acad. Sci. USA 88, 3392-3396.
15. 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., Barr,P.J., Miyamura,T., McHutchinson,J. and Houghton,M. (1992) Proc. Natl. Acad. Sci. USA 89, 3468-3472.
16. Gunji,T., Kato,N., Mori,S., Ootsuyama,Y., Hijikata,M., Imawari,M. and Shimotohno,K. (1992) Int. J. Cancer, in press.