*Journal of Viral Hepatitis*, 1995, 2, 235–241

# Sequence analysis of the hepatitis C virus (HCV) core gene suggests the core protein as an appropriate target for HCV vaccine strategies

Y. Hitomi,[1] W. M. McDonnell,[1,2] A. A. Killeen[3] and F. K. Askari[1] [1]*Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Medical Centre, Ann Arbor, MI,* [2]*Veterans Administration Medical Center, Ann Arbor, MI, and* [3]*Department of Pathology, University of Michigan Medical Centre, Ann Arbor, MI, USA*

*Received 20 March 1995; accepted for publication 13 June 1995*

SUMMARY. Hepatitis C virus (HCV) is a major healthproblem with a prevalence of 1% in the United States population, and a significant percentage of infected patients progress to chronic liver disease and cirrhosis. Interferon therapy has demonstrated that the immune system can be modulated to alter the acute course of the disease, but long-term treatments remain elusive. Prevention of hepatitis C infection is therefore an important strategy to mitigate the impact of this disease. Initial attempts at vaccination have focused on recombinant envelope vaccines, which have shown an ability to protect against very low titre challenges of HCV in chimps. The need for vaccines capable of protecting against higher titre challenges has led to the search for alternative vaccine strategies. The most highly conserved structural protein in the HCV genome is the core protein, and vaccine strategies targeting the core protein have been proposed to increase vaccine efficacy. The variability of HCV core sequences and genotypes in the Ann Arbor patient population are not known, and the present study was undertaken to assess the theoretical feasibility of developing a HCV core vaccine by excluding promiscuous core (C) gene variability as a mechanism of vaccine failure. Results of nucleotide and deduced amino acid sequence analysis from 13 of 14 patients studied reveal a 93% nucleotide and 96.4% amino acid core sequence homology in the C gene regions studied. Genotype analysis revealed four of 14 to be type 1a and nine of 14 to be type 1b with one infection not being sufficiently characterized to determine genotype. These results demonstrate a sufficiently high degree of conservation of HCV core sequences in our patient population to permit design of a vaccine directed against core protein.

*Keywords:* chronic hepatitis, cytotoxic immunity, nucleoprotein, viral hepatitis.

## INTRODUCTION

In 1989 the hepatitis C virus (HCV) was identified as the major causitive agent of transfusion associated non-A, non-B (NANB) hepatitis. Older data studying NANB hepatitis suggested that approximately 50% of infected people became chronic carriers and of those 20% would eventually develop cirrhosis. Newer information on HCV is more alarming. At least 70–80% of people infected with HCV become chronic carriers, and the number who develop significant liver disease may be higher than previously suspected [1]. The rate of spontaneous viral clearance is extremely low [2]. Efforts to treat chronic infection with medical interventions such as steroids or interferon have not been curative. Aggressive hepatic transplantation has failed to cure the disease as hepatic grafts become reinfected with HCV. Serum immune globulins that have been effective in preventing hepatitis A and hepatitis B infections are not effective prophylaxis against HCV following needle stick accidents.

Abbreviations: C gene, HCV core gene; HCV, hepatitis C virus; NANB, non A, non B hepatitis; RT-PCR, reverse transcription-polymerase chain reaction.

Correspondence: Dr Frederick K. Askari, University of Michigan Medical Center, 5520 MSRB I, 1150 West Medical Center Drive, Ann Arbor, MI, 48109-0680, USA.

One strategy for combatting this common infection is prophylactic vaccination. Developing a vaccine against HCV is challenging because antibodies against HCV have not been protective and envelope proteins are hypervariable. Initial attempts at HCV vaccination have focused on recombinant envelope protein vaccination strategies. In studies by Chiron Corporation (Emmeryville, CA), vaccines have shown a protective response to 10 Chimp Infectious Units, but not to higher titre viral loads. Several investigators have expressed concerns about the inability of anti-HCV envelope antibodies to elicit a more effective protective response [3,4].

Vaccine strategies that focus on the HCV core protein are attractive for two reasons:

1 the core protein is the most conserved HCV structural protein, and
2 results with vaccines against influenza A core protein, another virus with variable envelope proteins, suggest vaccine primed cytotoxic responses to viral core proteins might be protective in HCV [5].

Influenza A nucleoprotein DNA immunization has been shown to protect immunized mice from a lethal dose of influenza A [6]. Protection existed across several strains of influenza with variable envelope proteins. Given these encouraging results with influenza core protein vaccines, efforts are underway to develop a HCV vaccine using the C gene product.

In this context, the present study was undertaken to analyse more fully the sequence of the HCV C gene in viral isolates derived from HCV infected patients at the University of Michigan Medical Center in Ann Arbor. Numerous studies have reported the C gene sequence in single isolates, or in isolates derived from patients worldwide representing the 14 different genotypes [7–9]. Given the numerous genotypes and subtypes of HCV, it may prove necessary to develop multiple vaccines to prevent infection with all known strains. If a broad range of C gene sequences existed in isolates from a given area, this might provide a mechanism for vaccine failure. The present study was undertaken to analyse the C gene variability in isolates derived from the geographic area served by the University of Michigan Medical Center in Ann Arbor, Michigan. Analysis of C gene sequence variability in our area was of interest to assess whether a core protein vaccine strategy might be effective in preventing HCV infection in our population and to determine which HCV genotype would be appropriate to use in vaccine development. By expanding the pool of C gene sequence data and characterizing the genotypes in our isolates these data may also provide further insight into the epidemiology and biology of the HCV virus.

## MATERIALS AND METHODS

### *Serum and liver samples*

Serum and liver samples were obtained from patients cared for at the University of Michigan Medical Center in Ann Arbor, Michigan. At the time of hepatic transplantation, a liver sample was obtained from a patient known to be infected by HCV based on HCV recombinant immunoblot assay (RIBA) and liver pathology demonstrating peri-portal inflammation. Serum samples for HCV core sequence analysis were obtained from 14 consecutive HCV RIBA and HCV RNA positive patients whose serum was referred to the University of Michigan Molecular Diagnostics Lab for HCV RNA assay by reverse transcription-polymerase chain reaction (RT-PCR) amplification of the HCV 5' untranslated region.

### *Generation of first strand cDNA and PCR amplification*

RNA was extracted from liver and serum by phenol extraction and alcohol precipitation (Trizol, GIBCO BRL, Grand Island, NY) [10]. Single-stranded DNA was generated by *in vitro* transcription with M-MLV reverse transcriptase (GIBCO BRL, Grand Island, NY) using antisense primers as specified in Fig. 1. PCR was performed for 40 cycles in a thermal cycler (Perkin Elmer Cetus, Norwalk, CT). Reaction cycles were optimized to generate reaction product under the following conditions: denaturation at 94°C for 30 s; annealing at 45°C for 30 s; and, elongation at 72°C for 2 min. The reaction mixture contained 12.5 microlitres of first-strand cDNA solution, 10 mM Tris (pH 7.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 0.2 mM dNTP, 1 mM of each primer, and 1.25 U of *Taq* DNA polymerase (Boehringer Mannheim; Indianapolis, IN). All reactions were performed in duplicate with negative controls from serum not infected with HCV.

### *Sequencing reactions*

PCR products were electrophoresed on 5% acrylamide

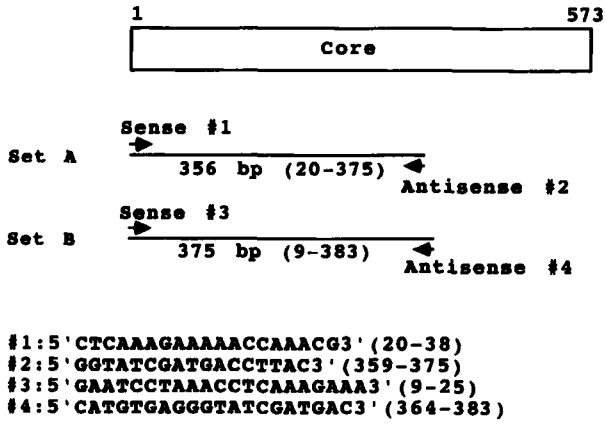


**Fig. 1** Amplification strategy. The two sets of oligonucleotide primers used for PCR amplification of HCV core sequences are shown above. A and B. Arrows indicate the position of sense and antisense primers. Nucleotides are numbered from first nucleotide of the open reading frame. The two sets of primers were deduced from HCV H77 [20].

gel and DNA bands of the predicted size (Fig. 2) were eluted. Aliquots of 10 to 40 ng of purified PCR products were directly sequenced by *Taq* DNA polymerase sequencing as described [11].

### *RIBA*

RIBA assays were performed by the Mayo Medical Laboratories (Rochester, MN, USA) and results were compared to the PCR data obtained in this study.

## RESULTS

Portions of the C gene from HCV isolates were successfully reverse transcribed and amplified by PCR from 13 of 14 patients confirmed to be infected with HCV by RIBA and RT-PCR (sequences deposited in GenBank, accession number pending). Interspersed control samples were always PCR negative for HCV RNA in over a dozen additional samples. PCR products were directly sequenced in duplicate. C gene sequence analysis of the Ann Arbor isolates revealed that no two isolates from different patients were identical, suggesting that the data reported here reflects authentic HCV sequence and not contamination.

The first screening of samples was performed with PCR primer set A (Fig. 1) which generated an anticipated 356 bp fragment in six of 14 samples as seen in Fig. 2(a). As the goal was to compare C gene sequences in a high percentage sample of our HCV patient population rather than generate a large non-specific database, a second set of primers was designed to the C gene sequence by comparing isolate AA 34 to the prototype HCV H77 sequence and designing primers to homologous sequences. All negative samples from the first set of screening amplifications were then analysed using a second set of primers and seven of eight of these reactions generated a 375 bp fragment of the anticipated size as seen in Fig. 2(b). Amplified C gene fragments were therefore available for sequence analysis from 93% of the isolates that were initially examined.

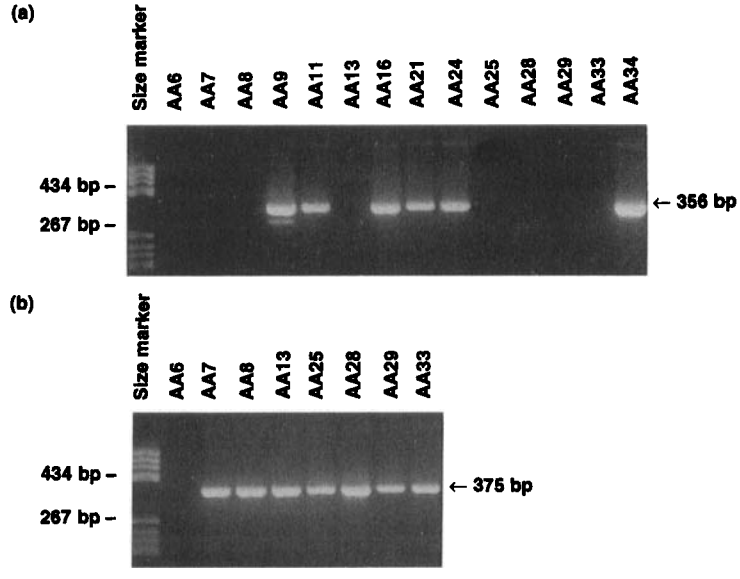


**Fig. 2** Agarose gel characterization of PCR products. HCV core amplification by RT-PCR using set A (a) and B (b) primers. RNA samples from serum or liver of HCV infected patients were transcribed into single-stranded cDNA and then amplified by 40 cycles of PCR. Aliquots of PCR products were subjected to 2.5% agarose gel electrophoresis and stained with ethidium bromide. Six samples out of fourteen are positive with set A primers (a). Seven of eight samples, negative with primer set A, are positive using primer set B (b).

**Table 1** Sequence homology of HCV isolates to prototype

| Isolate | % identity | | |
|---|---|---|---|
| homology* | Amino acid homology† | Nucleotide | |
| AA7- 1 | 97.4 | 96.7 | |
| 2 | 97.4 | 96.7 | |
| AA8- 1 | 98.2 | 98.9 | |
| 2 | 98.2 | 98.9 | |
| AA13- | 1 | 98.9 | 100 |
| 2 | 98.9 | 100 | |
| AA29- | 1 | 97.8 | 97.8 |
| 2 | 97.8 | 97.8 | |
| AA9- 1 | 90.1 | 95.6 | |
| 2 | 90.1 | 95.6 | |
| AA11- | 1 | 90.8 | 96.7 |
| 2 | 90.8 | 96.7 | |
| AA16- | 1 | 92.7 | 97.8 |
| 2 | 92.7 | 97.8 | |
| AA21- | 1 | 90.1 | 95.6 |
| 2 | 90.1 | 95.6 | |
| AA24- | 1 | 91.6 | 96.7 |
| 2 | 91.6 | 96.7 | |
| AA25- | 1 | 89.7 | 94.5 |
| 2 | 89.7 | 94.5 | |
| AA28- | 1 | 90.8 | 94.5 |
| 2 | 90.8 | 94.5 | |
| AA33- | 1 | 90.5 | 94.5 |
| 2 | 90.5 | 94.5 | |
| AA34- | 1 | 90.1 | 94.5 |
| 2 | 90.5 | 94.5 | |

* Nucleotide homology was calculated by comparing nucleotide sequence 58 to 330.
† Amino acid homology was calculated by comparing amino acid sequence 20 to 110.

Following sequencing reactions for both strands of the amplified C gene HCV isolates, the nucleotide and amino acid sequences of each isolate were determined for each reaction and compared with the prototype HCV-1 as reported in Table 1. Nucleotide homology was calculated by comparing nucleotide sequence from nt 58 to nt 330 of the putative HCV C gene for 26 amplifications from 13 unique isolates to those previously published for HCV-1 [12,13]. Nucleotide homology ranged from 89.7% to 97.4% with a mean homology of 93%. Amino acid homology to HCV-1 was calculated by comparing deduced amino acid sequence from amino acids 20 to 110 of the putative C protein sequence. Amino acid homology ranged from 94.5% to 100% with a mean homology of 96.4% in comparison with HCV-1.

It was interesting to determine which genotypes of HCV infect the patient population at the University of Michigan in Ann Arbor. The sequence of nt 186 to nt 221 was compared from each isolate to the known C gene sequences of the various HCV genotypes [9]. By making this comparison, it was possible to assign four isolates to genotype 1a and nine isolates to genotype 1b. All sequenced isolates were therefore assigned to either genotype 1a or genotype 1b, and only one sample which was C gene PCR negative could not be assigned to a specific genotype (Fig. 3).

One important issue was conservation of amino acid sequences at known HCV core T-cell recognition epitopes. For each isolate, the amino acid sequence from aa 41 to aa 49 was compared with a known HCV T-cell epitope previously defined in HCV-1, as summarized in figure 4 [14]. Ten of the thirteen isolates had an identical sequence from aa 41 to aa 49 to the prototype HCV-1, and three isolates had a single amino acid threonine to proline substitution at aa 49. Comparing the amino acid sequence for each isolate with a second T-cell epitope defined for HCV-TK from amino acids 81 to 100, one isolate matched the HCV-TK sequence [15]. Eight isolates had a single amino acid substitution at aa 91, four from leucine to cysteine, and four from leucine to methionine. Four other isolates had a tyrosine to phenylalanine substitution at aa 81 and a leucine to methionine substitution at aa 91.

The deduced amino acid sequences obtained from the sequenced core region were analysed with regard to amino acid composition and putative functional regions [8]. The predicted amino acid sequence of the Ann Arbor isolates also showed a high content of proline residues and conservation of 10 arginine and lysine residues invariant from position 39–62 with the exception of isolate AA-28, in which there was an arginine to serine (R to S) substitution at position 61. There was a cluster of five conserved tryptophan residues from position 76–106. With the exception of the position 61 amino acid substitution noted in isolate AA-28, two putative nuclear localization signals were conserved: one PRRGPR from amino acid 38–42 and the second PRGRRQP from position 58–64. The Ann Arbor isolates also demonstrated a putative SPRG DNA-binding motif at position 99–102 with the exception of isolate AA-29 which contained a R to S substitution at position 101.

**Table 2** Correlation of detection of antibodies to HCV (RIBA) with PCR results

| | PCR | | | | | RIBA |
|---|---|---|---|---|---|---|
| Isolate | 5' | Core | c22* | c33† | c100† | 5-1-1† |
| AA6 | + | – | ++++ | +/– | – | – |
| AA7 | + | + | ++++ | ++++ | +++ | ++ |
| AA8 | + | + | ++++ | ++++ | ++++ | ++++ |
| AA9 | + | + | ++++ | + | +/– | – |
| AA11 | + | + | ++++ | ++++ | ++ | ++ |
| AA13 | + | + | ++++ | ++++ | +++ | – |
| AA16 | + | + | ++++ | ++++ | – | – |
| AA21 | + | + | ++++ | ++++ | ++++ | ++++ |
| AA24 | + | + | ++++ | ++++ | +++ | ++ |
| AA25 | + | + | ++++ | ++++ | – | – |
| AA28 | + | + | ++++ | ++++ | +++ | +++ |
| AA29 | + | + | ++++ | ++++ | ++++ | ++++ |
| AA33 | + | + | ++++ | ++++ | +/– | – |
| AA34 | + | + | NA | NA | NA | NA |

* A recombinant core antigen of HCV.
† Recombinant antigens of HCV non-structural region.
NA, not available.

Finally, the sequence results were compared to RIBA results for the patients whose serum was subjected to analysis for the HCV C gene, as summarized in Table 2. Comparing the 12 sequenced serum isolates, all were positive for the c22 core antigen. Antibodies to non-structural proteins were detected in all samples, although AA 9 showed a weak signal to c33 and c100. The other sequenced isolates showed a strong signal for c33, and seven of these showed a strong signal for c100 and 5-1-1 as well. It was interesting to note that the only isolate which was PCR negative for the C gene, AA 6, did have a positive RIBA for c22 but not c33, c100 or 5-1-1. Antibodies were detected to the core protein but not the non-structural proteins.

## DISCUSSION

The ideal HCV vaccine would confer lifelong protective immunity against all known isolates of the virus with no deleterious side-effects. Given the hypervariability of envelope proteins, one would anticipate difficulty in protecting against numerous HCV strains with an envelope vaccine. Construction of a HCV vaccine presents similar challenges to those posed in constructing an influenza vaccine. Recent reports of protective immunization against influenza using an expressed core protein immunization strategy have raised hopes that a similar strategy may be applied to HCV prophylaxis.

Prior to commiting to a long-term effort to develop a HCV vaccine directed at the core protein, it was necessary to define C gene isolates in our patient population so a vaccine could be designed with appropriate C gene products. The present study represents the most extensive survey to date of HCV core sequences reported in patients from a single geographic region [8,9,12].

HCV RIBA and PCR of the 5' untranslated region are known to be more sensitive than PCR of the C gene region. The study design calls for screening patient samples for HCV first by RIBA and then by PCR 5' region primers prior to attempting HCV C gene PCR on sequential samples referred to the Molecular

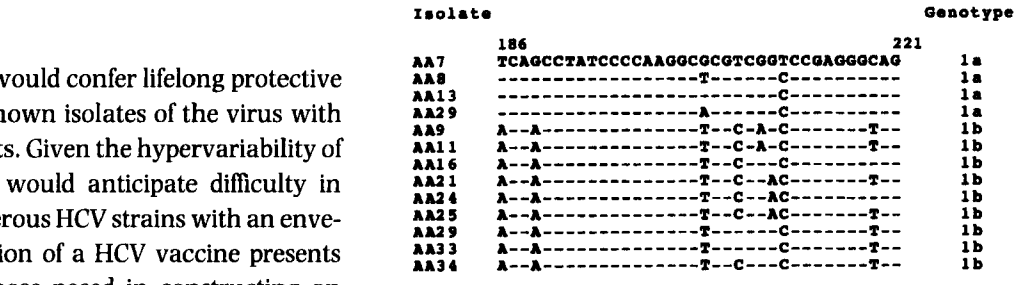


Fig. 3 HCV genotypes of HCV Ann Arbor isolates. The HCV genotype of each isolate was deduced by comparing genotype specific sequences with those previously published [9]. HCV genotypes were assigned according to the standardized nomenclature [17].

| Isolate | HCV-1 | HCV-TK |
|---|---|---|
| | 41 49 | 81 100 |
| | GPRLGVRAT | YPWPLYGNEGLGWAGWLLS |
| AA7 | --------- | ----------C-------- |
| AA8 | --------- | ----------C-------- |
| AA13 | --------- | ----------C-------- |
| AA29 | --------- | ----------C-------- |
| AA9 | --------- | ----------M-------- |
| AA11 | --------- | ----------M-------- |
| AA16 | --------- | ---------- |
| AA21 | --------- | ----------M-------- |
| AA24 | --------- | ----------M-------- |
| AA25 | ---------P | F----------M-------- |
| AA28 | --------- | F----------M-------- |
| AA33 | ---------P | F----------M-------- |
| AA34 | ---------P | F----------M-------- |

**Fig. 4** Conservation of amino acid sequences at HCV core protein cytotoxic T-cell epitopes. Deduced amino acid sequences of 13 clones were compared with two previously defined core epitopes [14,15].

Diagnostics Lab for analysis. Using the first two assays to define infection with HCV, C gene PCR was successful for 13 of 14 patients studied. The sequences presented here therefore reflect the C gene sequences from 93% of the patients enrolled in the study. By using two sets of primers and adjusting the reaction conditions, the sensitivity of C gene PCR was optimized.

Nucleotide identity was defined in comparison with an arbitrary prototype, HCV-1. The range of homology from 89.7% to 97.4% was consistent with the variability seen in genotypes 1a and 1b in a study of isolates from a worldwide patient sample [8]. Mutations in nucleotide sequence which led to amino acid changes were observed throughout the C gene sequence analysed comparing one isolate with another. Infection with two different strains of virus was not observed in any of the duplicate isolates analysed.

Since at least 14 genotypes of HCV have been reported around the world it was interesting to determine which genotypes of HCV were seen in our patients. All isolates were classified to genotype 1a or 1b which is consistent with genotypes observed in other samples from this continent [14,15]. No unanticipated genotypes were isolated [16,17]. The consistency of HCV genotypes infecting our patient population denotes a degree of consistency or stability in the HCV gene pool within our region. Such conservation of core sequence homology could prove critical to the success of a new vaccine.

Different individuals may generate cytotoxic responses to different regions of the C protein, presumably in part due to variability of MHC proteins. Epitope studies have used different C protein sequences to stimulate cytotoxic assays further complicating the definition of cytotoxic T-cell epitopes. One of the advantages of using an expressed core protein vaccine strategy would be generation of cytotoxic responses, so amino acid sequences from our isolates were compared with two putative T-cell epitopes. In the first case, a HCV-1 epitope was compared with our isolates with a single amino acid substitution at the end of the epitope identified in three of 13 isolates. Variation in a second T-cell epitope defined by aa 81 to 100 was more pronounced, with 12 of 13 isolates demonstrating substitutions at least at aa 91. These variable amino acid substitutions in the centre of this putative T-cell epitope might well effect cytotoxic responses to that epitope, although what role, if any, such mutations might play in vaccine failure is unclear [18–20]. Clearly, careful attention will have to be devoted to such issues as vaccine trials progress in chimps and ultimately humans.

The amino acid sequence data lend further support to hypothesized putative functional roles for several regions of the core protein. The presence of nuclear localization and DNA binding motifs lend further support to the hypothesis that the C protein of HCV might also function as a gene regulatory protein [21]. The isolates in this study have highly conserved features which support the previously stated hypothesized roles of different regions of the C protein [8]. As further C protein studies are performed, the postulated roles of given regions will either be further confirmed or refuted.

The C gene sequence of one patient, AA 6, was not amplified by PCR. Failed C gene PCR correlated with the weakest antibody responses to c33, c100 and 5-1-1 of the studied isolates. The weak RIBA response could be for one of two reasons caused by either low viral titre or protein mutation substantial enough to shift the cross-reactivity of protein standards to generated HCV antibodies. The presence of an antibody response to c22 infers that core sequences were present at one time and the positive signal in the PCR 5' region primers implies the ongoing presence of HCV virus. The failure of C gene PCR may reflect base pair mismatches to both sets of primers or an inadequate viral titre for successful reaction.

In summary, analysis of C gene sequences isolated from a small sample of our HCV-infected patient population confirmed substantial conservation of C gene sequences. Such conservation of C gene sequences is critical to the success of a HCV vaccine directed against the C gene product. All sequenced isolates were classified to genotype 1a or 1b, consistent with genotypes isolated from other regions of this continent and excluding

broad genotype heterogeneity as a reason for vaccine failure in our patient population. In addition to recent advances in gene-transfer vaccine development, these data provide further cause for optimism that a HCV core protein vaccine will successfully be developed.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the kind and generous advice of the late Mark Roth MD in the early development of this study; his enthusiasm carried forward to its completion. We would like to acknowledge support from the Department of Internal Medicine, Division of Gastroenterology and University of Michigan Office of Vice President for Research, Rackham Graduate School, as well as generous supplemental equipment access from Drs Rob Todd, David Ginsburg and Jim Wilson. This work was supported in part by the University of Michigan Medical Center NIH Cancer Center Pilot Feasibility Award, and Glaxo Gastroenterology Award.

## REFERENCES

1 Takahasi M, Yamada G, Miyamoto R, Doi T, Endo H, Tsuji T. Natural course of chronic hepatitis C. *Am J Gastroenterol* 1993; 88: 240–243.
2 Alter MJ, Margolis HS, Krawczynski K *et al.* The natural history of community acquired hepatitis C in the United States. *N Engl J Med* 1992; 327: 1899–1905.
3 Choo QL, Kuo G, Ralston R *et al.* Vaccination of chimpanzees against infection by the hepatitis C virus. *Proc Natl Acad Sci USA* 1994; 91: 1294–1298.
4 Alter MJ. The detection, transmission, and outcome of HCV infection. *Infect Agents Dis* 1993; 2: 155–166.
5 McDonnell WM. Gene transfer as a new mode of vaccination: implications for HCV. *Hepatology* 1993; 18: 694–698.
6 Ulmer JB, Donnelly JJ, Parker SE, Rhodes GH, Felgner PL, Liu MA. Heterologous protection against influenza by injection of DNA encoding a viral protein. *Science* 1993; 259: 1745–1749.
7 Takamizawa A, Mori C, Fuke I *et al.* Structure and organization of the hepatitis C virus Genome isolated from human carriers. *J Virol* 1991; 65(3): 1105–1113.
8 Bukh J, Purcell RH, Miller RH. Sequence analysis of the core gene of 14 hepatitis C virus genotypes. *Proc Natl Acad Sci USA* 1994; 91: 8239–8243.
9 Cha TA, Beall E, Irvine B *et al.* At least five related, but distinct, hepatitis C viral genotypes exist. *Proc Natl Acad Sci USA* 1992; 89: 7144–7148.
10 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate phenol-chloroform extraction. *Anal Biochem* 1987; 162: 156–159.
11 Lee JS. Alternative dideoxy sequencing of double-stranded DNA by cyclic reactions using Taq polymerase. *DNA Cell Biol* 1991; 10(1): 67–73.
12 Choo QL, Richman KH, Han JH *et al.* Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci USA* 1991; 88: 2451–2455.
13 Choo QL, Kuo G, Weiner A *et al.* Identification of the major, parenteal non-A non-B hepatitis agent (hepatitis C virus) using a recombinant approach. *Semin Liver Dis* 1992; 12: 279–288.
14 Koziel MJ, Dudley D, Afdhal N *et al.* Hepatitis C virus specific cytotoxic T lymphocytes recognize epitopes in the core and envelope proteins of HCV. *J Virol* 1993; 67: 7522–7532.
15 Kita H, Moriyama T, Kaneko T *et al.* HLA B44 restricted cytotoxic T lymphocytes recognizing an epitope on hepatitis C virus nucleocapsid protein. *Hepatology* 1993; 18: 1039–1044.
16 Bukh J, Purcell RH, Miller RH. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proc Natl Acad Sci USA* 1993; 90: 8234–8238.
17 Simmonds P, Alberti A, Alter HJ *et al.* A proposed system for the nomenclature of hepatitis C viral genotypes. *Hepatology* 1994; 19: 1321–1324.
18 Shimizu YK, Hijikata M, Iwamoto Alter HJ, Purcell RH, Yoshikura H. Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. *J Virol* 1994; 68: 1494–1500.
19 Houghton M, Weiner A, Han J, Kuo G, Choo QL. Molecular biology of the hepatitis C viruses: implications for diagnosis and control of viral disease. *Hepatology* 1991; 14: 381–388.
20 Ogata N, Alter HJ, Miller RH, Purcell RH. Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc Natl Acad Sci USA* 1991; 88: 3392–3396.
21 Shih CM, Lo SJ, Miyamura T, Chen SY, Lee YHW. Suppression of hepatitis B virus expression and replication by hepatitis C virus core protein in HUH 7 cells. *J Virol* 1993; 67: 5823–5832.