Genetic Diversity of Hepatitis B Virus as an Important Factor Associated with Differences in Clinical Outcomes

Yasuhito Tanaka and Masashi Mizokami

Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan

(See the article by Livingston et al., on pages 5–11.)

Approximately 350 million people worldwide are persistently infected with hepatitis B virus (HBV) [1], which has been classified into 8 genotypes (A–H) by means of molecular evolutionary analyses [2–4]. HBV genotypes have different geographic distributions [5]. Genotype D is ubiquitous and scattered worldwide, whereas genotype A is prevalent in sub-Saharan Africa, North America, and Europe and genotypes B and C are common in Asia. Genotype E is mainly restricted to western Africa, and genotype F is considered to be indigenous to aboriginal populations in Central and South America [6]. In addition, a number of genotype F–specific phylogenetic clusters have been attributed to local populations in different geographic areas [7, 8], such as Argentina [9], Venezuela [10], and Central America [11], as well as in the drainage area of the Yukon and Kuskokwim Rivers in southwestern Alaska [12]. Genotype F also has been detected sporadically in Brazil [13], North America [6, 14], Europe [8], and Polynesia [15]. Genotype H has been found in Mexico and Central America [4, 6]. Genotype G has been detected infrequently, and its epidemiologic profile is unclear.

Genotypes are further subdivided into subgenotypes, on the basis of phylogenetic relationships [16]. Subgenotype Aa/A1 has been identified in South Africa and Asia, whereas subgenotype Ae/A2 has been detected in Europe and the United States [17, 18]. Subgenotype Ac/A3 has been detected in central and western Africa [19]. Sugauchi et al. [20] identified 2 subgenotypes within genotype B in Asian countries. One subgenotype (Bj/B1) is the authentic genotype B and is indigenous in Japan, whereas another subgenotype (Ba/B2) is predominant in Asian countries other than Japan and exhibits a recombination with genotype C over the precore region and core gene [20, 21]. Recently, subgenotypes have also been recognized in genotypes C and D [16, 22–24].

There is increasing evidence regarding the influence of HBV genotypes/subgenotypes on liver disease, in both acute and chronic HBV infections [5, 25]. Owing to the geographic distribution of genotype prevalence, comparative analyses have been restricted mainly to the predominant genotypes—namely, genotype B versus genotype C in some Asian countries and genotype A versus genotype D in Europe and India [26–28]—although a few multinational studies also have compared >2 genotypes [29, 30]. In Asian cohorts, genotype C has been associated with a higher frequency of cirrhosis or hepatocellular carcinoma (HCC) and a weaker patient response to interferon-α–based treatment, compared with genotype B [31]. Similarly, genotype-related differences have been reported for long-term outcomes of chronic infection with HBV genotype A, D, or F [32]: rates of sustained biochemical remission and clearance of HBV DNA and hepatitis B surface antigen were significantly higher among patients infected with genotype A than among patients infected with genotype D or F. In addition, the frequency of death related to liver disease has been reported to be higher for genotype F than for genotype A or D [32].

In this issue of the *Journal*, Livingston et al. [12] report that genotype F strains were found to be significantly associated with the occurrence of HCC among Alaska Native people, compared with genotype A, B, C, or D. In a multiple logistic regression analysis of region-, sex-, and birth-adjusted cohorts, this association was still significant ($P < .001$; odds ratio, 8.9 [95% confidence interval, 4.4–17.8]). In addition, Alaska Native people infected with genotype F were younger at the time of diagnosis of HCC: the median age at diagnosis of HCC was lower for patients

Received 12 September 2006; accepted 13 September 2006; electronically published 21 November 2006.

Potential conflicts of interest: none reported.

Financial support: Ministry of Health, Labour and Welfare of Japan (grant-in-aid H16-kanen-3); Ministry of Education, Culture, Sports, Science and Technology of Japan (grant-in-aid for Scientific Research [C] [18590741]); Toyoaki Foundation.

Reprints or correspondence: Dr. Yasuhito Tanaka, Dept. of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Kawasumi, Mizuho, Nagoya 467-8601, Japan (ytanaka@med.nagoya-cu.ac.jp).

The Journal of Infectious Diseases 2007; 195:1–4

0022-1899/2007/19501-0001$15.00

infected with genotype F than for patients infected with other genotypes (22.5 vs. 60 years, respectively; $P = .002$). These novel findings should be of interest to clinicians and researchers, especially in regions where genotype F predominates.

Differences in age at diagnosis of HCC between genotypes/subgenotypes have also been reported in Asia: younger patients in Taiwan were more likely to be infected with genotype B (Ba) than C, and subgenotype Ba was associated with the development of HCC in young patients who did not have cirrhosis [33]. Similarly, in South Africa, a 4.5-fold increased risk for HCC and a younger age at diagnosis (by 6 years) was reported among HBV carriers infected with genotype A (Aa), compared with those infected with non-A genotypes [34].

Mutations of the HBV genome in the basic core promoter (BCP) region (T1762/A1764) and a stop codon–generating mutation in the precore region (A1896) have been found to be associated with HBeAg seroconversion and with viral replication [35]. Both the BCP and precore stop codon mutations are often found in patients with advanced liver disease, such as HCC [36–38]. Beyond these mutations, a C→T mutation in the upstream regulatory sequence (C1653T) has been associated with fulminant hepatitis [39], possibly because of its location in the α box, which is a strong activating element of both enhancer II and core promoter [40]. Takahashi et al. [41] reported that the C1653T and T1753V (not T) mutations were associated with the progression of liver disease from chronic hepatitis to cirrhosis and/or HCC, and further studies have indicated that the T1653 mutation is found frequently in patients with HCC and genotype C infection [42, 43]. Associations for these mutations are being actively investigated within the context of different genotypes, which could suggest different mechanisms for genotypes A, D, and C and even their subgenotypes [42–44]. Virologic differences between genotypes and subgenotypes have been confirmed by recent in vitro studies [45].

In Costa Rica, where HBV genotype F is the most prevalent (96%), a BCP double mutation was detected in carriers with chronic infection (41.4%) but not in patients who recovered from acute infection [46]. Livingston et al. [12] showed no association between the BCP mutation and HCC among Alaska Native people infected with genotype F; in Alaska, the prevalence of the BCP mutation among patients with HCC and genotype F infection (41%), versus that in control patients with genotype F infection but without HCC (70% including mixture), was higher than that previously reported for blood donors with genotype F infection in Argentina (33%) [9], suggesting possible differences in genotype F between different population groups. On the other hand, for genotypes A, C, and D, a significantly higher frequency of BCP mutations was observed among patients with HCC, versus that among the genotype-matched control patients. Thus, this relationship between BCP mutations and HCC may be valid for some HBV genotypes but not for genotype F. Different mechanisms and pathways might play a role in different genotypes, and this issue is far from being completely understood.

A clear association between genotype F and the 1896A precore stop codon mutation was found among Argentinean blood donors (36/48 [75%]; $P < .05$) [9]. This association was described in a previous report on HBV-infected patients of Hispanic origin in Central America [47], where most genotype F strains contain T1858. Among Alaska Native people, however, a relatively low prevalence of the precore mutation and a similar frequency of the mutation in patients with HCC (22%) and control patients (40% including mixture) were found [12]. Differences in the viral sequences of HBV genotype F in patients with HCC and those without HCC would be important to investigate, as would the search for the genetic features of genotype F that are potentially associated with carcinogenesis.

The risk of development of chronic HBV infection is inversely related to age at time of infection, that is, to the transmission route. The risk is highest when the infection is acquired perinatally (vertical transmission) or in early childhood and decreases with age [48, 49]. This is likely to be associated with different host immunological responses to HBV infection at different ages. Different transmission patterns predominate in different countries, and little is known regarding whether these differences are associated with the specific behavior of different genotypes, particularly genotypes F and D in Alaska.

Environmental influences could also play an important role in disease outcome. Although an association with aflatoxin B_1, previously found to be associated with HCC in Africa, was not found in a study of traditional food in Alaska [50], another study has reported that at least 13% of Alaska Native people with HCC appear to have chronic hepatitis not related to HBV or hepatitis C virus infection, suggesting the possibility of some form of unrecognized chronic liver disease predisposing to HCC [51]. Additional studies are needed to investigate the possible overlap of such factors among young patients with HCC who are infected with genotype F. Furthermore, to help develop more-effective therapies, the focus of hepatitis B research should concentrate on viral, host, and environmental factors that determine clinical outcomes in patients with chronic HBV infection.

References

1. Lee WM. Hepatitis B virus infection. N Engl J Med **1997**; 337:1733–45.
2. Norder H, Courouce AM, Magnius LO. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. Virology **1994**; 198:489–503.
3. Stuyver L, De Gendt S, Van Geyt C, et al. A new genotype of hepatitis B virus: complete

genome and phylogenetic relatedness. J Gen Virol **2000**; 81:67–74.
4. Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol **2002**; 83:2059–73.
5. Miyakawa Y, Mizokami M. Classifying hepatitis B virus genotypes. Intervirology **2003**; 46: 329–38.
6. Norder H, Hammas B, Lee SD, et al. Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. J Gen Virol **1993**; 74:1341–8.
7. Mbayed VA, Barbini L, Lopez JL, Campos RH. Phylogenetic analysis of the hepatitis B virus (HBV) genotype F including Argentine isolates. Arch Virol **2001**; 146:1803–10.
8. Norder H, Arauz-Ruiz P, Blitz L, Pujol FH, Echevarria JM, Magnius LO. The T(1858) variant predisposing to the precore stop mutation correlates with one of two major genotype F hepatitis B virus clades. J Gen Virol **2003**; 84: 2083–7.
9. Franca PH, Gonzalez JE, Munne MS, et al. Strong association between genotype F and hepatitis B virus (HBV) e antigen–negative variants among HBV-infected Argentinean blood donors. J Clin Microbiol **2004**; 42: 5015–21.
10. Blitz L, Pujol FH, Swenson PD, et al. Antigenic diversity of hepatitis B virus strains of genotype F in Amerindians and other population groups from Venezuela. J Clin Microbiol **1998**; 36:648–51.
11. Arauz-Ruiz P, Norder H, Visona KA, Magnius LO. Molecular epidemiology of hepatitis B virus in Central America reflected in the genetic variability of the small S gene. J Infect Dis **1997**; 176:851–8.
12. Livingston SE, Simonetti JP, McMahon BJ, et al. Hepatitis B virus genotypes in Alaska Native people with hepatocellular carcinoma: preponderance of genotype F. J Infect Dis **2007**; 195:5–11 (in this issue).
13. Naumann H, Schaefer S, Yoshida CF, Gaspar AM, Repp R, Gerlich WH. Identification of a new hepatitis B virus (HBV) genotype from Brazil that expresses HBV surface antigen subtype adw4. J Gen Virol **1993**; 74:1627–32.
14. Chu CJ, Keeffe EB, Han SH, et al. Prevalence of HBV precore/core promoter variants in the United States. Hepatology **2003**; 38:619–28.
15. Couroucé AM, Plancon A, Soulier JP. Distribution of HBsAg subtypes in the world. Vox Sang **1983**; 44:197–211.
16. Norder H, Courouce AM, Coursaget P, et al. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. Intervirology **2004**; 47: 289–309.
17. Kramvis A, Weitzmann L, Owiredu WK, Kew MC. Analysis of the complete genome of subgroup A′ hepatitis B virus isolates from South Africa. J Gen Virol **2002**; 83:835–9.
18. Sugauchi F, Kumada H, Acharya SA, et al. Epidemiological and sequence differences between two subtypes (Ae and Aa) of hepatitis B virus genotype A. J Gen Virol **2004**; 85: 811–20.
19. Kurbanov F, Tanaka Y, Fujiwara K, et al. A new subtype (subgenotype) Ac (A3) of hepatitis B virus and recombination between genotypes A and E in Cameroon. J Gen Virol **2005**; 86:2047–56.
20. Sugauchi F, Orito E, Ichida T, et al. Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. J Virol **2002**; 76: 5985–92.
21. Sugauchi F, Orito E, Ichida T, et al. Epidemiologic and virologic characteristics of hepatitis B virus genotype B having the recombination with genotype C. Gastroenterology **2003**; 124:925–32.
22. Huy TT, Ushijima H, Quang VX, et al. Genotype C of hepatitis B virus can be classified into at least two subgroups. J Gen Virol **2004**; 85:283–92.
23. Kramvis A, Kew M, Francois G. Hepatitis B virus genotypes. Vaccine **2005**; 23:2409–23.
24. Tanaka Y, Orito E, Yuen MF, et al. Two subtypes (subgenotypes) of hepatitis B virus genotype C: a novel subtyping assay based on restriction fragment length polymorphism. Hepatol Res **2005**; 33:216–24.
25. Kramvis A, Kew MC. Relationship of genotypes of hepatitis B virus to mutations, disease progression and response to antiviral therapy. J Viral Hepat **2005**; 12:456–64.
26. Erhardt A, Blondin D, Hauck K, et al. Response to interferon alfa is hepatitis B virus genotype dependent: genotype A is more sensitive to interferon than genotype D. Gut **2005**; 54:1009–13.
27. Orito E, Mizokami M, Sakugawa H, et al. A case-control study for clinical and molecular biological differences between hepatitis B viruses of genotypes B and C. Japan HBV Genotype Research Group. Hepatology **2001**; 33: 218–23.
28. Thakur V, Guptan RC, Kazim SN, Malhotra V, Sarin SK. Profile, spectrum and significance of HBV genotypes in chronic liver disease patients in the Indian subcontinent. J Gastroenterol Hepatol **2002**; 17:165–70.
29. Westland C, Delaney W, Yang H, et al. Hepatitis B virus genotypes and virologic response in 694 patients in phase III studies of adefovir dipivoxil. Gastroenterology **2003**; 125: 107–16.
30. Janssen HL, van Zonneveld M, Senturk H, et al. Pegylated interferon alfa-2b alone or in combination with lamivudine for HBeAg-positive chronic hepatitis B: a randomised trial. Lancet **2005**; 365:123–9.
31. Kao JH, Wu NH, Chen PJ, Lai MY, Chen DS. Hepatitis B genotypes and the response to interferon therapy. J Hepatol **2000**; 33: 998–1002.
32. Sanchez-Tapias JM, Costa J, Mas A, Bruguera M, Rodes J. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in Western patients. Gastroenterology **2002**; 123:1848–56.
33. Kao JH, Chen PJ, Lai MY, Chen DS. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. Gastroenterology **2000**; 118:554–9.
34. Kew MC, Kramvis A, Yu MC, Arakawa K, Hodkinson J. Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu-speaking sub-Saharan Africans. J Med Virol **2005**; 75:513–21.
35. Buckwold VE, Xu Z, Chen M, Yen TS, Ou JH. Effects of a naturally occurring mutation in the hepatitis B virus basal core promoter on precore gene expression and viral replication. J Virol **1996**; 70:5845–51.
36. Laskus T, Radkowski M, Nowicki M, Wang LF, Vargas H, Rakela J. Association between hepatitis B virus core promoter rearrangements and hepatocellular carcinoma. Biochem Biophys Res Commun **1998**; 244:812–4.
37. Baptista M, Kramvis A, Kew MC. High prevalence of 1762(T) 1764(A) mutations in the basic core promoter of hepatitis B virus isolated from black Africans with hepatocellular carcinoma compared with asymptomatic carriers. Hepatology **1999**; 29:946–53.
38. Kao JH. Hepatitis B virus genotypes and hepatocellular carcinoma in Taiwan. Intervirology **2003**; 46:400–7.
39. Ogata N, Miller RH, Ishak KG, Purcell RH. The complete nucleotide sequence of a precore mutant of hepatitis B virus implicated in fulminant hepatitis and its biological characterization in chimpanzees. Virology **1993**; 194: 263–76.
40. Yuh CH, Chang YL, Ting LP. Transcriptional regulation of precore and pregenomic RNAs of hepatitis B virus. J Virol **1992**; 66: 4073–84.
41. Takahashi K, Ohta Y, Kanai K, et al. Clinical implications of mutations C-to-T1653 and T-to-C/A/G1753 of hepatitis B virus genotype C genome in chronic liver disease. Arch Virol **1999**; 144:1299–308.
42. Ito K, Tanaka Y, Orito E, et al. T1653 mutation in the box alpha increases the risk of hepatocellular carcinoma in patients with chronic hepatitis B virus genotype C infection. Clin Infect Dis **2006**; 42:1–7.
43. Tanaka Y, Mukaide M, Orito E, et al. Specific mutations in enhancer II/core promoter of hepatitis B virus subgenotypes C1/C2 increase the risk of hepatocellular carcinoma. J Hepatol **2006**; 45:646–53.
44. Tanaka Y, Hasegawa I, Kato T, et al. A case-control study for differences among hepatitis B virus infections of genotypes A (subtypes Aa and Ae) and D. Hepatology **2004**; 40:747–55.
45. Sugiyama M, Tanaka Y, Kato T, et al. Influence of hepatitis B virus genotypes on the intra- and extracellular expression of viral DNA and antigens. Hepatology **2006**; 44:915–24.
46. Leon B, Taylor L, Vargas M, et al. HBx M130K and V131I (T-A) mutations in HBV genotype

F during a follow-up study in chronic carriers. Virol J **2005**; 2:60.
47. Arauz-Ruiz P, Norder H, Visona KA, Magnius LO. Genotype F prevails in HBV infected patients of Hispanic origin in Central America and may carry the precore stop mutant. J Med Virol **1997**; 51:305–12.
48. Hyams KC. Risks of chronicity following acute hepatitis B virus infection: a review. Clin Infect Dis **1995**; 20:992–1000.
49. McMahon BJ, Alward WL, Hall DB, et al. Acute hepatitis B virus infection: relation of age to the clinical expression of disease and subsequent development of the carrier state. J Infect Dis **1985**; 151:599–603.
50. De Benedetti VM, Welsh JA, Trivers GE, et al. p53 is not mutated in hepatocellular carcinomas from Alaska Natives. Cancer Epidemiol Biomarkers Prev **1995**; 4:79–82.
51. Chung YH, Di Bisceglie AM, McMahon BJ, et al. Hepatocellular carcinoma not related to hepatitis B virus infection among Alaska Natives. Int J Circumpolar Health **1999**; 58: 208–13.

Available online at www.sciencedirect.com
ScienceDirect
Virology 361 (2007) 27–33
VIROLOGY
www.elsevier.com/locate/yviro

A reduction in selective immune pressure during the course of chronic hepatitis C correlates with diminished biochemical evidence of hepatic inflammation

Kousuke Hanada [a], Yasuhito Tanaka [b,*], Masashi Mizokami [b], Takashi Gojobori [a], Harvey J. Alter [c]

[a] National Institute of Genetics, Yata 1111, Mishima, Shizuoka, Japan
[b] Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Kawasumi, Mizuho, Nagoya 467-8601, Japan
[c] National Institutes of Health, Clinical Center Department of Transfusion Medicine, Bethesda, MD 20892, USA

Received 22 June 2006; returned to author for revision 20 September 2006; accepted 27 October 2006
Available online 15 December 2006

Abstract

It is considered that selection pressure exerted by the host immune response during early HCV infection might influence the outcome of that infection particularly as it relates to persistence or clearance of the agent. However, it is unclear whether positive selection pressure plays a role in determining the severity of hepatitis C during the course of persistent HCV infection. To address the evolutionary mechanism by which HCV escapes from the host immune response and to assess the relationship between viral evolution and hepatic inflammation, we determined 57 sequences (3–5 serial samples per patient) from 5 individuals with persistent HCV infection of genotype 1a who were under long-term follow-up ranging from 15.6 to 21.6 years. We applied a novel method to estimate serial alternations of selective pressure against the HCV enveloped region and compared this to fluctuation in transaminase level over time. Positive selection pressure was reduced over time postinfection, as evidenced by a reduction in nonsynonymous substitutions in the later phase of infection. Furthermore, serum transaminase, as a measure of inflammatory necrosis of hepatocytes, was reduced in parallel with decreased positive selection pressure. These results suggest that during persistent HCV infection, the virus faces diminished immune pressure over time, either from mutation to an immune resistant sequence or from immunologic exhaustion, and that this diminished immune attack is reflected in diminished inflammatory activity. This observation may be applicable to other viruses characterized by a slow rate of disease progression.

Keywords: HCV; RNA virus; Evolution; Positive selection; Immune response; Selective pressure; Persistent infection; E2; Alanine aminotransferase (ALT)

Introduction

Hepatitis C virus (HCV) was identified in 1989 as the major etiologic agent of non-A and non-B hepatitis (Choo et al., 1989). As many as 170 million people worldwide are infected with HCV and it is the world's leading cause of chronic liver disease (WHO, 1997). Once HCV infects a human, some patients are able to spontaneously clear the infection, but the vast majority develops persistent infection. It is postulated that persistent infection is predicted, at least in part, by an impaired HCV-specific immune response that can be measured directly by studies of T-cell immunity (Chang et al., 1997), and indirectly by the degree of immune pressure exerted against viral envelope proteins (Franco et al., 1995). In contrast, a strong immune response and high level of immune pressure early in HCV infection can result in viral clearance and a self-limited infection (Klenerman et al., 2000; Missale et al., 2004). Thus, the vigor of positive selection pressure mediated through the early HCV-specific immune response is thought to be a determining factor. However, the vigor of positive selection pressure has been unclear in persistent HCV infection.

Persistent HCV infection is usually clinically mild, but 20% of patients progress to severe chronic hepatitis and cirrhosis, occasionally culminating in hepatocellular carcinoma (HCC)

* Corresponding author. Fax: +81 52 842 0021.
E-mail address: ytanaka@med.nagoya-cu.ac.jp (Y. Tanaka).

0042-6822/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2006.10.041

over the course of two or more decades (Kiyosawa et al., 1991). The precise reasons for this very slow rate of disease progression in chronic hepatitis C remain an enigma, but since liver damage is thought to be immune-mediated the implication is that an impaired HCV-specific immune response simultaneously allows for viral persistence and insidious liver damage (Ehl et al., 1998). Although the immune response is impaired, it generates sufficient immune pressure to generate nonsynonymous viral escape mutations in acute infection (Farci et al., 2000). Thus a delicate balance seems to be achieved wherein viral variants survive the immune response while maintaining essential replicative functions and liver damage is minimized by an impaired immune attack. The present study addresses how HCV escapes the host immune response and how this might translate to persistent, but low-level liver damage and an insidious course that is measured in decades rather than years.

A common way of measuring selective pressure is determining the ratio of the number of synonymous substitutions per synonymous site (Ds) to the number of nonsynonymous substitutions per nonsynonymous site (Dn). Generally, this ratio has been studied in the E2 envelope protein region of the HCV genome (Farci et al., 1994; Sheridan et al., 2004; Shimizu et al., 1994). Much evidence supports that this region is the primary positively selected region in which Dn is larger than Ds. To estimate the serial alternation of selective pressure of E2 over time, we applied a variation of previous methods (Chen and Wang, 2005; Christie et al., 1999; Manzin et al., 2000). We first demonstrated the evolutionary relationship among time points by the phylogenetic tree, and then estimated the selective pressure by the Dn/Ds ratio over the time from the initial infection. In addition, the alanine aminotransferase (ALT) level was employed as an indirect serial measure of hepatocyte injury as previously reported (Matsuoka et al., 1994; Pradat et al., 2002).

Using these combined methods, we succeeded in determining the serial alternation of both selective pressure and damage to hepatocytes throughout the long course of chronic HCV infection.

Results

Comparison of selective pressure among 5 patients

Taking account of the difference in years of isolation for each clone, we constructed the phylogenetic trees in five patients (Fig. 1). To know the variation of the selection pressure among the five patients, the mean Dn/Ds ratio was determined in each patient. There was a four-fold variation of Dn/Ds ratio among patients (Fig. 2). Moreover, Dn/Ds ratio was positively correlated ($r=0.76$) with the mean ALT values (29–73 time points), though the association was not significant because of

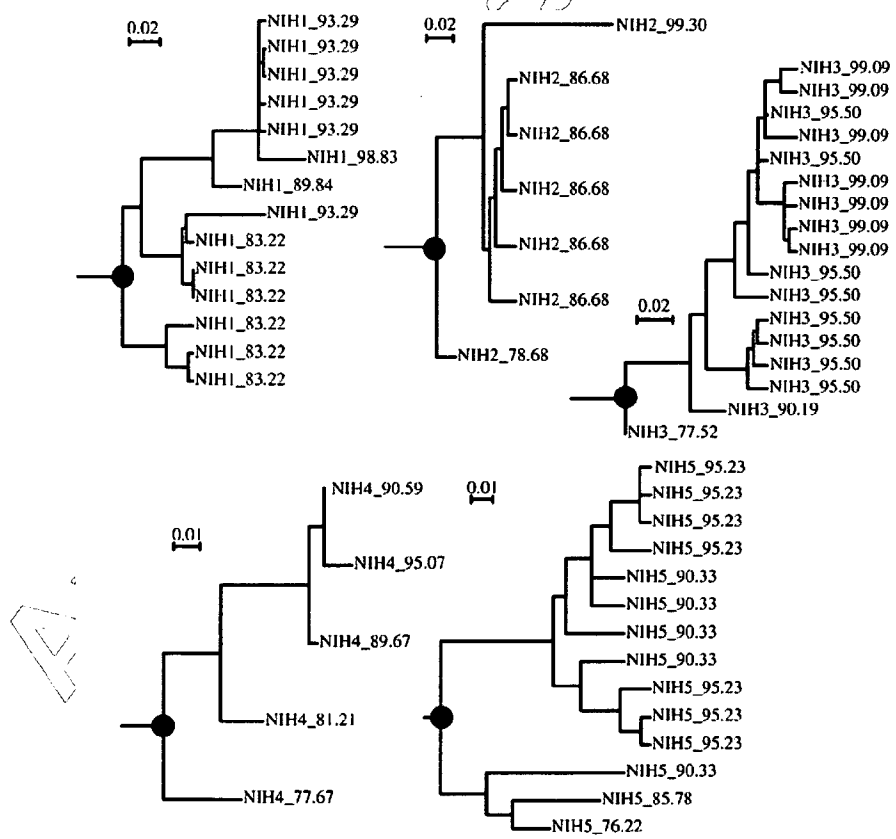

Fig. 1. Phylogenetic tree assuming the molecular clock. A phylogenetic tree represents the viral evolutionary process occurring in an infected patient. The postfix of a sequence represents the year at which the sequence was isolated. In the isolated years, 1900 was subtracted from the dominical year. Solid circle indicates the location of most recent common ancestor (MRCA) in each patient.

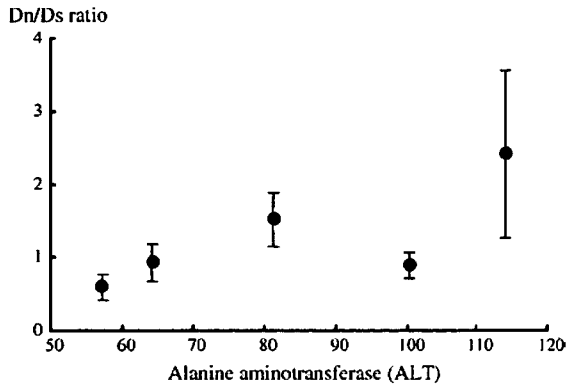

Fig. 2. The relationship between Dn/Ds ratio and ALT value among 5 patients. Each dot indicates a single patient. *X*-axis is the mean of ALT values over 15 years postinfection. *Y*-axis is the mean Dn/Ds ratio representing the selective pressure. Standard deviation of Dn/Ds ratio in each patient is estimated by the bootstrap method.

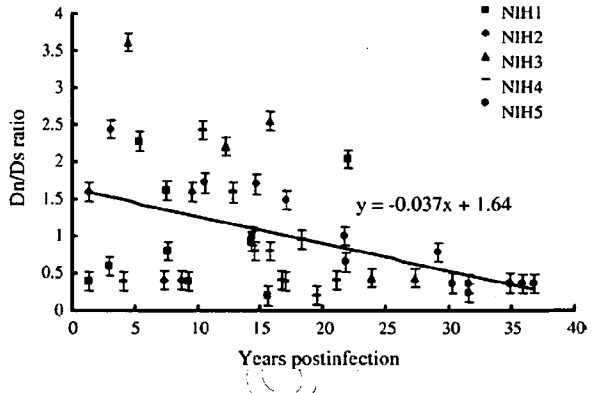

Fig. 3. The relationship between Dn/Ds ratio and years postinfection. Each dot indicates a branch in the phylogenetic tree. *X*-axis is the mean of the two past yeas in each branch. *Y*-axis is Dn/Ds ratio. Standard deviation of Dn/Ds ratio is shown in each branch. The regression line shows the relationship between Dn/Ds ratio and years postinfection.

the small number of patients (Fig. 2). Of note, histological data also indicated that patients with a high Dn/Ds ratio tended to have higher HAI inflammatory grades at the latest points available (NIH1:1993, NIH2:1999, NIH3:1999, NIH4:1995 and NIH5:1995), except for the one patient who had bridging fibrosis (Table 1 and Fig. 2).

Serial changes in selective pressure over the duration of infection

The overall association between Dn/Ds ratio and the duration of infection was examined in the five patients (Fig. 3). The association was negatively correlated ($r=-0.44$, $P<0.01$). Although it is usually considered that nonsynonymous substitutions are continuous in order to maintain escape from immune pressure, our results showed that HCV kept reducing nonsynonymous substitutions over time in comparison with synonymous substitutions. This apparent decrease in selection pressure paralleled a decrease in hepatocellular injury as reflected in declining ALT levels (Fig. 4). Thus, there was a concomitant decline in Dn/Ds ratio and ALT with increasing time after initial infection.

Adaptations in genomic regions during persistent infection

The regions affected by positive selection pressure were identified by the comparison between the number of synonymous and nonsynonymous substitutions within the hypervariable region of the HCV envelope. Two regions showed higher Dn/Ds ratios in persistently infected patients (Fig. 5). The first epitope is in the hyper-variable region and has been reported to be a B-cell epitope (Weiner et al., 1992). The second epitope corresponds to amino acid residues 457–462, a region whose function has not been defined. These two regions appear to have been positively selected in response to immune pressure but they are not thought to be functionally constrained regions because the Dn/Ds ratio of the two regions is much higher than 1.

Two previous papers found positively selected regions in HCV by the comparison between the rate of synonymous substitution and the rate of nonsynonymous substitutions (Suzuki and Gojobori, 2001; Sheridan et al., 2004). However vaccine design using only synonymous and nonsynonymous substitutions leads a dilemma that HCV could escape from the vaccine-induced immune system by nonsynonymous substitutions leading to escape mutants even if the vaccine were constructed against these critical epitopes. Therefore, we examined chemically constrained codon regions in the two regions (Fig. 6). When the proportion of chemically radical changes in a window (4 codon sites) is smaller than the other windows, the window is considered to be a more chemically constrained region. Such constrained regions could be the targets because the potential amino acid replacements are limited in their ability to undergo chemical change. Three

Table 1
Clinical data of 5 study subjects with chronic hepatitis C virus

Patients	Sample interval (year)	Sample numbers	Age	Gender	Race[a]	ALT values mean±s.e.[b] (Numbers of sample collection)	Degree of chronic hepatitis[c]
1	15.6	14	71	M	B	100.3±23.5 (71)	10/3
2	20.6	7	55	M	W	81.1±10.1 (66)	5/1
3	21.6	17	52	F	B	57.2±12.4 (29)	4/1
4	17.7	5	84	M	W	113.9±38.0 (74)	10/1
5	19	14	76	M	W	64.2±8.4 (41)	5/1

[a] B: Black African American, W: White Caucasian.
[b] s.e.: standard error.
[c] The scores for grading/staging were based on HAI and Knodell fibrosis score (Knodell et al., 1981).

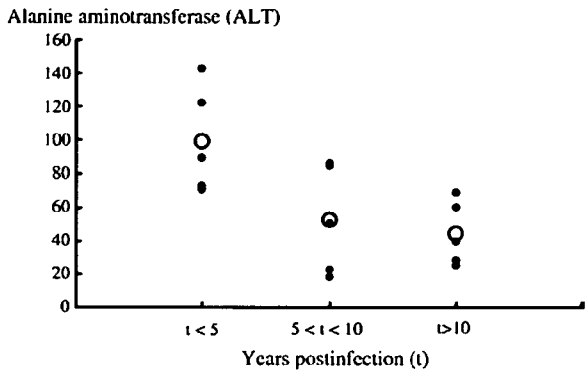

Fig. 4. The relationship between ALT level and years postinfection. Small solid circles indicate the means of ALT values during the years postinfection for each patient (0–5, 5–10, >10). Large open circles indicate the mean ALT for each time period.

windows (residues 406–409, 407–410 and 411–414) were identified in which radical changes were significantly less than the changes observed in all other windows ($P<0.05$).

Discussion

Our analyses found a positive correlation between immune selection pressure and hepatocyte damage over the duration of HCV infection. Hence the vigor of the immune response may drive both the extent of ALT elevations and the degree of viral complexity and diversity. In this study, we used serum ALT levels to reflect sequential changes in liver cell injury since serial liver biopsies could not be justified. While this is an indirect measure of inflammation-associated injury, there was good correlation between mean ALT level and the Knodell HAI score in the single liver biopsy obtained for each patient. It has been previously established that high ALT levels correlate with more severe outcomes in HCV infection and that persistently normal or low-level ALT elevations are generally associated with mild histologic findings. However, severe histologic lesions are sometimes observed even in the face of normal serum ALT. (Matsuoka et al., 1994; Pradat et al., 2002). From these observations, we believe that the overall disturbance of hepatocytes can be estimated by mean ALT values during follow-up and that correlations between positive selection pressure and liver injury can be validly estimated by serial measurement of serum ALT levels.

The main finding of the present study is that both the extent of positive selection pressure and the degree of liver injury were reduced sequentially during the long course of chronic HCV infection. We believe these findings to be interrelated since the vigor of the immune response drives both liver injury and viral escape. In this regard, several studies have reported that the immune response against HCV weakens in the later phases of infection (Takaki et al., 2000; Wang et al., 2002; Wedemeyer et al., 2002). These results suggest that during persistent HCV infection, the virus faces diminished immune pressure over time, either from mutation to an immune resistant sequence or from immunologic exhaustion, and that this diminished immune attack is reflected in diminished inflammatory activity. This would in part explain the indolent course of chronic HCV infection and the absence of spontaneous resolution once chronicity is established.

Although anti-viral therapy has become increasingly effective, the prime goal for the control of global HCV infection is the development of a prophylactic vaccine and a secondary goal would be development of a therapeutic vaccine that could enhance response to antiviral therapy by boosting the immune system. To examine the potentially immunogenic target, we utilized a modification of the method of Suzuki and Gojobori (Suzuki and Gojobori, 1999) in which a phylogenetic tree was constructed for each patient and the number of synonymous and nonsynonymous substitutions estimated throughout the phylogenetic tree. Using this method, a B-cell epitope within the hyper-variable region of the HCV envelope was identified as being under immune pressure and hence this region might serve as an important immunogen for vaccine development The

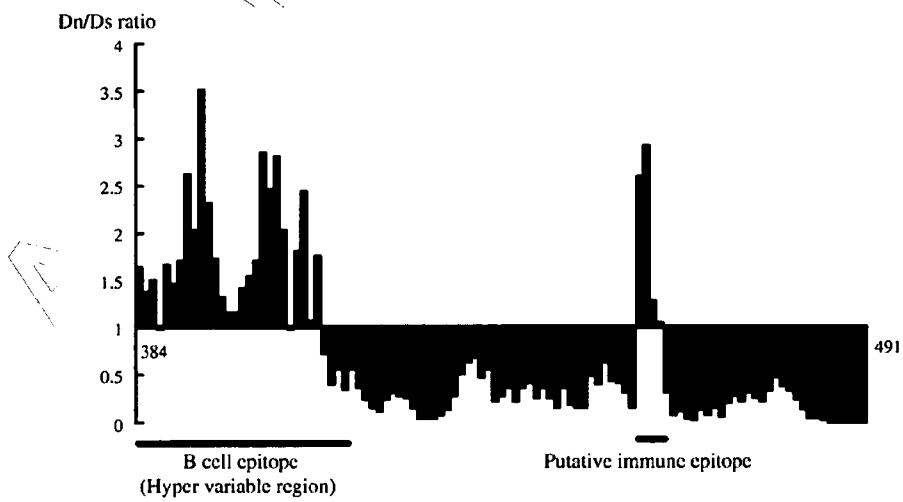

Fig. 5. Dn/Ds ratio by the window analysis (window size=4 codon sites). *X*-axis is the starting codon site for the window analysis. *Y*-axis is Dn/Ds ratio for each window.

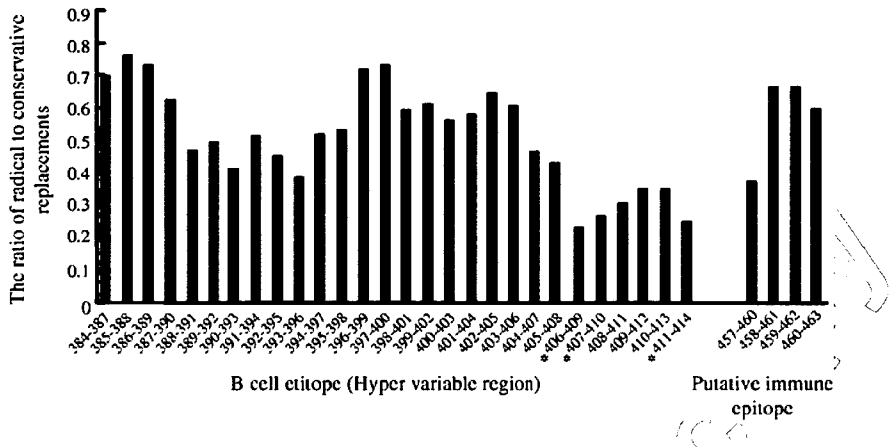

Fig. 6. Proportion of chemically radical replacements in two immune related epitopes. *X*-axis is the starting and ending codon site for the window analysis. *Y*-axis is the ratio of chemically radical to conservative replacements for each window. Significant codon sites were indicated as the asterisk [chemically radical changes are smaller than conservative changes ($P<0.05$)].

candidate epitopes in this study were in the E2 (envelope) region of the HCV genome. Similar approaches could be applied to the entire HCV genome to potentially define other critical epitopes.

Although this study demonstrated that both positive selection pressure and liver damage were reduced over the time in those with persistent infection and mild chronic hepatitis, a much larger population with diverse hepatitis outcomes would need to be studied by phylogenetic analysis to ascertain if the viral evolutionary process, as a reflection of host immune response, truly dictates the outcome of chronic liver disease in persistent HCV infection.

Materials and methods

Long-term serial samples

We determined fifty-seven sequences (varying from 3 to 5 serial samples in each patient) from 5 patients with HCV genotype 1a. All samples were obtained from patients enrolled in prospective studies of transfusion-associated hepatitis conducted at National Institutes of Health since 1972 (Alter et al., 1989). The interval between the first and last sample in the 5 patients for whom serial samples were available ranged from 15.6 to 21.6 years. The clinical data for each patient are shown in Table 1. ALT values shown in Table 1 are the mean of ALT values obtained over the duration of follow-up, ranging up to 15 years. All patients had chronic hepatitis without cirrhosis, and 3 of 5 patients had only mild inflammation. Available histological results in the five patients were indicated by the histological activity index (HAI), also known as the Knodell score (Knodell et al., 1981).

RNA extraction, PCR, and sequencing

To extract total RNA from human sera, a RNA isolator (Genosys, The Woodlands, TX) was used, with slight modification from the manufacturer's instructions; Superscript II RNase H⁻ reverse transcriptase (GIBCO/BRL) was used for reverse transcription. To reduce the number of artifactual substitutions occurring during PCR, Platinum Pfx DNA Polymerase (GIBCO/BRL) with very high fidelity was used. For the PCR in the E2 region, long PCR fragments (core to NS2 regions) were amplified by HCV1- and 2-core-s1 (346–373): GCACGAATCCTAAACCTCAAAGAAAAAC and HCV1ab-ns2-as (2540–2565): CTATCAGCAGCATCATCCACAAG-CAG. The first products were nested by using the following primers: 45AT-s1 (1290–1310): CGCATGGCNTGGGAYAT-GATG [n=any, Y=C or T] and HCVE2-as2 (1848–1868): GAAGCAATAYACYGGRCCACA. The PCR products of a single target molecule were directly sequenced. If mixed sequences were detected by direct sequencing, the HCV amplified products were ligated into pCR-Blunt II-TOPO Vector and used to transform DH5α high-efficiency competent cells according to the manufacturer's protocol (Invitrogen Corp.). The plasmid DNA was purified using the QIAprep Spin Miniprep kit (Qiagen Inc, Hilden, Germany) and the presence of the inserts confirmed by digestion with *Eco*R1. Sequencing was performed on more than 5 clones per each point. All clones were sequenced with Prism Big Dye (Applied Biosystems, Foster City, California, USA) in an ABI 3100 DNA automated sequencer. The determined sequences have accession numbers (AB272166–AB272225).

Estimation of selective pressure in phylogenetic tree

Multiple alignment in each patient was made to match the coding regions by the computer program, clustalW (Thompson et al., 1994). The consensus sequence of all the sequences was used as an outgroup to locate the position of the root in each phylogenetic tree. The phylogenetic tree was constructed by the maximum likelihood method premised on a molecular clock in Fig. 1 (Rambaut, 2000). Taking into account the difference in years of isolation among the sequences, we estimated the divergence time of all nodes in the tree. We then inferred ancestral nucleotide sequences at all nodes of the phylogenetic tree by the maximum likelihood method (Yang et al., 1995). These analyses were conducted by the program called PAML.

The number of substitutions per synonymous or nonsynonymous site (Ds and Dn) was independently estimated for all branches by the modified Nei–Gojobori method (Zhang et al., 1998).

We estimated two kinds of selective pressures, namely the selective pressure for each patient and the serial selective pressure over the years since the onset of infection. To determine the mean selective pressure for each patient, the mean Dn/Ds ratio was simply estimated in all the branches for each patient. To determine the serial selective pressure for the years since the onset of infection, the infected year for each patient was estimated as the divergence time of the most recent common ancestor (MRCA) in all the nodes of the phylogenetic tree (Fig. 1). The year postinfection of each node was defined as the divergence time of the node minus the divergence time of MRCA. The year postinfection of each terminal node was defined as the isolation time of the terminal branch minus the divergence time of MRCA. This indicated that 2 years could be applied in each branch. The mean year of 2 years was used as the year postinfection in each branch. In order to see the change of selective pressure during the past year, we examined both Dn/Ds ratio and the mean time postinfection in each branch. When synonymous substitutions do not occur in several short branches, the Dn/Ds ratio cannot be calculated. Therefore, to equalize the weight for the two kinds of substitution, we used only the branches that included both synonymous and nonsynonymous substitutions.

Estimation of positively selected regions

Positively selected pressures of a particular region were identified by the modified version of Suzuki and Gojobori method (Suzuki and Gojobori, 1999). In this method, a phylogenetic tree was reconstructed for each patient by the maximum likelihood method. The ancestral sequence was inferred at every node in the phylogenetic tree using the maximum likelihood method. Then, the numbers of synonymous and nonsynonymous substitutions throughout the phylogenetic tree were estimated in each branch for each codon site. We independently summed up the total numbers of synonymous (Ns) and nonsynonymous (Nn) substitutions occurring in 5 patients for each codon site. The mean numbers of synonymous (Cs) and nonsynonymous (Cn) sites were calculated for each codon site by the modified Nei–Gojobori method. Ds and Dn was calculated as Ns/Cs and Nn/Cn as p-distance, respectively. To see positively selected pressures, Dn/Ds was calculated by the window analysis. The window size used was 4 codon sites.

Moreover, we examined the degree of chemical bias in E2 region. All amino acids were classified into 6 groups by Zhang's classification, which took account of polarity and volume (Zhang, 2000). When an amino acid replacement occurred between the groups, the amino acid replacement was defined as the chemically radical change. Analogously, when an amino acid replacement occurred within the single group, the amino acid replacement was defined as the chemically conservative change. The number of chemically radical and conservative replacements was counted throughout the phylogenetic tree in each codon site. The total numbers of radical changes and conservative changes were calculated in each window (the window size=4 codon sites). The ratio of radical to conservative changes was estimated in each window. *Z*-value of each ratio was estimated from the average and standard deviation in all the windows. *P*-value was calculated from *Z*-value based on the assumption that the distribution of ratios is the normal distribution.

Acknowledgments

We appreciate Thomas Nagylaki and two reviewers for useful comments. The work was supported by a grant-in-aid from the Ministry of Health, Labour and Welfare of Japan (H16-kanen-3), grants-in-aid for Young Scientists (A) from the Ministry of Education, Culture, Science, Sports of Japan (16689016) and Japan Science and Technology Agency.

References

Alter, H.J., Purcell, R.H., Shih, J.W., Melpolder, J.C., Houghton, M., Choo, Q.L., Kuo, G., 1989. Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. N. Engl. J. Med. 321, 1494–1500.

Chang, K.M., Rehermann, B., McHutchison, J.G., Pasquinelli, C., Southwood, S., Sette, A., Chisari, F.V., 1997. Immunological significance of cytotoxic T lymphocyte epitope variants in patients chronically infected by the hepatitis C virus. J. Clin. Invest. 100, 2376–2385.

Chen, S., Wang, Y.H., 2005. Multigene tracking of quasispecies in viral persistence and clearance of hepatitis C virus. World J. Gastroenterol. 11, 2874–2884.

Choo, Q.L., Kuo, G., Weiner, A.J., Overby, L.R., Bradley, D.W., Houghton, M., 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244, 359–362.

Christie, J.M., Chapel, H., Chapman, R.W., Rosenberg, W.M., 1999. Immune selection and genetic sequence variation in core and envelope regions of hepatitis C virus. Hepatology 30, 1037–1044.

Ehl, S., Klenerman, P., Zinkernagel, R.M., Bocharov, G., 1998. The impact of variation in the number of CD8(+) T-cell precursors on the outcome of virus infection. Cell. Immunol. 189, 67–73.

Farci, P., Alter, H.J., Wong, D.C., Miller, R.H., Govindarajan, S., Engle, R., Shapiro, M., Purcell, R.H., 1994. Prevention of hepatitis C virus infection in chimpanzees after antibody-mediated in vitro neutralization. Proc. Natl. Acad. Sci. U.S.A. 91, 7792–7796.

Farci, P., Shimoda, A., Coiana, A., Diaz, G., Peddis, G., Melpolder, J.C., Strazzera, A., Chien, D.Y., Munoz, S.J., Balestrieri, A., Purcell, R.H., Alter, H.J., 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science 288, 339–344.

Franco, A., Ferrari, C., Sette, A., Chisari, F.V., 1995. Viral mutations, T cell receptor antagonism and escape from the immune response. In: Sher, A., Ahmed, R. (Eds.), Current Opinion in Immunology. Current Biology Ltd., London, pp. 524–531.

Kiyosawa, K., Sodeyama, T., Tanaka, E., Nakano, Y., Furuta, S., Nishioka, K., Purcell, R.H., Alter, H.J., 1991. Hepatitis C in hospital employees with needlestick injuries. Ann. Intern. Med. 115, 367–369.

Klenerman, P., Lechner, F., Kantzanou, M., Ciurea, A., Hengartner, H., Zinkernagel, R., 2000. Viral escape and the failure of cellular immune responses. Science 289, 2003.

Knodell, R.G., Ishak, K.G., Black, W.C., Chen, T.S., Craig, R., Kaplowitz, N., Kiernan, T.W., Wollman, J., 1981. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology 1, 431–435.

Manzin, A., Solforosi, L., Debiaggi, M., Zara, F., Tanzi, E., Romano, L., Zanetti,

A.R., Clementi, M., 2000. Dominant role of host selective pressure in driving hepatitis C virus evolution in perinatal infection. J. Virol. 74, 4327–4334.

Matsuoka, S., Tatara, K., Hayabuchi, Y., Taguchi, Y., Mori, K., Honda, H., Itou, S., Yuasa, Y., Kuroda, Y., 1994. Serologic, virologic, and histologic characteristics of chronic phase hepatitis C virus disease in children infected by transfusion. Pediatrics 94, 919–922.

Missale, G., Cariani, E., Ferrari, C., 2004. Role of viral and host factors in HCV persistence: which lesson for therapeutic and preventive strategies? Dig. Liver Dis. 36, 703–711.

Pradat, P., Alberti, A., Poynard, T., Esteban, J.I., Weiland, O., Marcellin, P., Badalamenti, S., Trepo, C., 2002. Predictive value of ALT levels for histologic findings in chronic hepatitis C: a European collaborative study. Hepatology 36, 973–977.

Rambaut, A., 2000. Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16, 395–399.

Sheridan, I., Pybus, O.G., Holmes, E.C., Klenerman, P., 2004. High resolution phylogenetic analysis of Hepatitis C virus adaptation and its relationship to disease progression. J. Virol. 78, 3447–3454.

Shimizu, Y.K., Hijikata, M., Iwamoto, A., Alter, H.J., Purcell, R.H., Yoshikura, H., 1994. Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. J. Virol. 68, 1494–1500.

Suzuki, Y., Gojobori, T., 1999. A method for detecting positive selection at single amino acid sites. Mol. Biol. Evol. 16, 1315–1328.

Suzuki, Y., Gojobori, T., 2001. Positively selected amino acid sites in the entire coding region of hepatitis C virus subtype 1b. Gene 3 (276), 83–87.

Takaki, A., Wiese, M., Maertens, G., Depla, E., Seifert, U., Liebetrau, A., Miller, J.L., Manns, M.P., Rehermann, B., 2000. Cellular immune responses persist and humoral responses decrease two decades after recovery from a single-source outbreak of hepatitis C. Nat. Med. 6, 578–582.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Wang, H., Bian, T., Merrill, S.J., Eckels, D.D., 2002. equence variation in the gene encoding the nonstructural 3 protein of hepatitis C virus: evidence for immune selection. J. Mol. Evol. 54, 465–473.

Wedemeyer, H., He, X.S., Nascimbeni, M., Davis, A.R., Greenberg, H.B., Hoofnagle, J.H., Liang, T.J., Alter, H.J., Rehermann, B., 2002. Impaired effector function of hepatitis C virus-specific CD8+ T cells in chronic hepatitis C virus infection. J. Immunol. 169, 3447–3458.

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., Hutchinson, J.M., Houghton, M., 1992. Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: potential role in chronic HCV infections. Proc. Natl. Acad. Sci. U.S.A. 89, 3468–3472.

WHO, 1997. Hepatitis C: global update. Wkly. Epidemiol. Rec. 72, 341–344.

Yang, Z., Kumar, S., Nei, M., 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics 141, 1641–1650.

Zhang, J., 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. J. Mol. Evol. 50, 56–68.

Zhang, J., Rosenberg, H.F., Nei, M., 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc. Natl. Acad. Sci. U.S.A. 95, 3708–3713.