附件：封面格式

行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※※

臺灣地區B型肝炎病毒表面蛋白基因變異的盛行率及其臨床意義

※※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：■個別型計畫　□整合型計畫

計畫編號：NSC 90－2314－B002－182－MH

執行期間：90 年 8 月 1 日至 91 年 7 月 31 日

計畫主持人：劉俊人

共同主持人：高嘉宏

計畫參與人員：陳定信、陳培哲、賴明陽、邵文逸、張毓敏

本成果報告包括以下應繳交之附件：

□赴國外出差或研習心得報告一份

□赴大陸地區出差或研習心得報告一份

□出席國際學術會議心得報告及發表之論文各一份

□國際合作研究計畫國外研究報告書一份

執行單位：臺大醫院內科部

中 華 民 國 91 年 10 月 20 日

中文摘要

B 型肝炎病毒感染為一世界性問題，影響健康甚鉅，於臺灣更甚。慢性 B 型肝炎可能導致肝硬化，肝衰竭，和肝癌，但至今，急性感染後會導致慢性 B 型肝炎的免疫病理機制尚未完全明瞭。針對 B 型肝炎病毒表面蛋白的體液性免疫系統，是清除血液循環中病毒的主要機制，而細胞性免疫系統則有助於消滅細胞中的病毒顆粒。但相對的，若 B 型肝炎病毒表面蛋白基因發生變異，則病毒可能有機會逃避宿主免疫系統的攻擊而存活下來，而這些變異株也可能與慢性 B 型病毒感染後的臨床病程進展或持續感染有關。過去相關研究極為有限。臺灣地區以 *adw* 血清型/基因型 B-B 型肝炎病毒感染為最常見，這些 B 型肝炎病毒表面蛋白的 B 細胞及 T 細胞的免疫反應標的區是否也常常會出現變異，以及這些免疫反應標的區的病毒變異株對慢性 B 型肝炎病程有何影響，則尚未明瞭。

本研究探討在自然演變的情況下，不同血清型的 B 型肝炎病毒表面蛋白出現變異的情形。為研究這些病毒變異株，尤其是可以逃避宿主免疫力攻擊的病毒變異株，對慢性 B 型肝炎病程的影響，自感染初期便大規模長期追蹤病人是需要的，但此類臨床研究有時間上的限制。退而求其次，我們橫斷性地探討不同慢性 B 型肝炎病程病人體內 B 型肝炎病毒位於這些免疫反應標的區的變異盛行率，間接地藉以了解病毒變異株與疾病病程的相關性。

本研究共收集 122 位 B 型肝炎病毒感染患者，包括 33 位肝功能正常帶原者，31 位慢性肝炎患者，22 位肝硬化患者，及 36 位肝細胞癌患者。我們首先增幅並定序涵蓋 *a* 決定子(胺基酸 124 至 148)及第一型 MHC-限制性毒殺型 T 細胞的免疫反應標的區(胺基酸 28 至 51)的表面蛋白基因，所衍生胺基酸序列則進一步與病毒血清型 *adw* 或 *adr* 的共識胺基酸序列相對比，以找出胺基酸變異。最後以 GEE 合併 Poisson 模型分析單一序列單一胺基酸變異發生率(FEQ)與臨床或病毒特性的相關性。結果顯示整體 FEQ 是 1.21%，在肝細胞癌患者以及 50 歲以上患者則最高。相對照之下，9 個感染血清型 *adw* 且小於 50 歲的肝細胞癌患者有異常偏高的 FEQ。就血清型 *adw* 而言，胺基酸變異聚集於 *a* 決定子和 CTL 標的區，就血清型 *adr* 而言，變異較為分散，而且變異熱點在血清型 *adw* 和 *adr* 之間並不相同。

結論：B 型肝炎病毒表面蛋白 FEQ 與患者年齡和肝疾病嚴重度間有正相關，某些變異株可能造成病毒持續感染。

關鍵詞：血清型；B 型肝炎病毒；表面蛋白基因；B 細胞；T 細胞；免疫反應標的區；變異；盛行率

English abstract

Hepatitis B virus (HBV) infection is a major health problem worldwide, affecting approximately 350 million persons. The clinical consequences of HBV infection range from acute self-limited infection or fulminant hepatitis, to various forms of chronic infection, including asymptomatic carrier state, chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma (HCC). The humoral immune response to HBV surface protein is suggested to contribute to the clearance of circulating HBV particles, whereas the cellular immune response is responsible for the elimination of infected hepatocytes. Alternatively, the emergence of variant viruses that could escape from humoral or cellular immunity may be implicated in the persistence of HBV infection. However, the contribution of immune escape variants in the natural course of chronic HBV infection remains unknown.

The production of an antibody to the *a* determinant (amino acids 124 to 148), located within the major hydrophilic region of the surface gene, after vaccination usually protects against HBV infection. There is ample evidence that escape mutations within the *a* determinant occur under strong selective pressure by vaccination in infants born to HBsAg-positive mothers and by immunoprophylaxis in orthotopic liver transplantation recipients. In addition to the *a* determinant, the surface gene of HBV contains putative HLA class I-restricted cytotoxic T lymphocyte (CTL) epitope (amino acids 28 to 51). In general, viral mutations in CTL epitope could evade cellular immunity and thus contribute to persistent infection. However, the prevalence of escape mutations within the *a* determinant and CTL epitope in the natural course of chronic HBV infection was evaluated in only few studies.

To clarify the role of these immune escape mutants within HBV surface gene in the persistence or progress of chronic hepatitis B, long-term prospective follow-up of patients from the beginning of HBV infection is mandatory. However, according to the natural history of HBV infection, it will take more than 3 decades to complete the study. In this study, we alternatively evaluated the association of the frequency of amino acid variation within the immunogenic epitopes of surface gene with different disease stages of chronic HBV infection. The surface gene of HBV encompassing the *a* determinant (amino acids 124 to 148) and the putative HLA class I-restricted cytotoxic T lymphocyte (CTL) epitope (amino acids

28 to 51) were amplified and directly sequenced in 33 asymptomatic carriers (group I), 31 patients with chronic hepatitis (group II), 22 with cirrhosis (group III), and 36 with hepatocellular carcinoma (group IV). The amino acid sequences were subsequently compared with the consensus sequences of HBV serotype *adw* or *adr*. The frequency of amino acid variation per site per sequence (FEQ) was analyzed by generalized estimating equation with Poisson model after stratification by clinical and virological features. We found that the FEQ was 1.21% overall, and was highest in group IV patients and in patients above 50 years of age. In contrast, 9 group IV patients aged below 50 years who were infected with serotype *adw* had an inversely higher FEQ than those above 50; the age effect among hepatocellular carcinoma patients was significantly different from that among non-cancerous patients (P=0.04). Variation of amino acid clustered within *a* determinant and CTL epitope for serotype *adw* but was randomly distributed for serotype *adr*. Mutation hotspots differed between serotypes *adw* and *adr*. The FEQ of HBV surface protein is positively correlated with advancing age and severity of liver disease, and certain variants possibly contribute to the persistence of HBV infection.

Keywords : Serotype; hepatitis B virus; surface gene; B cell; T cell; immunogenic epitope; variation; prevalence

Attached please find our original article (**Liu CJ, Kao JH, Shau WY, Chen PJ, Chen DS. Naturally occurring hepatitis B virus surface gene variants: Correlated with viral serotypes and clinical stages of liver diseases. Journal of Medical Virology 2002;68:50-9**) disclosing the relevant sections of the projects.

Journal of Medical Virology 68:50–59 (2002)

Naturally Occurring Hepatitis B Surface Gene Variants in Chronic Hepatitis B Virus Infection: Correlation With Viral Serotypes and Clinical Stages of Liver Disease

Chun-Jen Liu,[1] Jia-Horng Kao,[1,2] Wen-Yi Shau,[2] Pei-Jer Chen,[1,2,3] Ming-Yang Lai,[1,2] and Ding-Shinn Chen[1,2]*

[1]*Department of Internal Medicine, National Taiwan University Hospital, Taiwan*
[2]*Graduate Institute of Clinical Medicine, National Taiwan University College of Medicine, Taiwan*
[3]*Hepatitis Research Center, National Taiwan University Hospital, Taiwan*

Virus variants escaping from host immunity may be implicated in the pathogenesis of hepatitis B virus (HBV) infection. In this cross-sectional study, the association was evaluated of the frequency of amino acid variation within the immunogenic epitopes of surface gene with different disease stages of chronic HBV infection. The surface gene of HBV encompassing the *a* determinant (amino acids 124–148) and the putative HLA class I restricted cytotoxic T lymphocyte (CTL) epitope (amino acids 28– 51) were amplified and directly sequenced in 33 asymptomatic carriers (Group I), 31 patients with chronic hepatitis (Group II), 22 with cirrhosis (Group III), and 36 with hepatocellular carcinoma (Group IV). The amino acid sequences were compared subsequently with the consensus sequences of HBV serotype *adw* or *adr*. The frequency of amino acid variation per site per sequence (FEQ) was analyzed by generalized estimating equation with Poisson model after stratification by clinical and virological features. The FEQ was 1.21% overall, and was highest in Group IV patients and in patients above 50 years of age. In contrast, nine Group IV patients aged below 50 years who were infected with serotype *adw* had an inversely higher FEQ than those above 50; the age effect among hepatocellular carcinoma patients was significantly different from that among non-cancerous patients ($P = 0.04$). Variation of amino acid clustered within *a* determinant and CTL epitope for serotype *adw* but was distributed at random for serotype *adr*. Mutation hotspots differed between serotypes *adw* and *adr*. The FEQ of HBV surface protein is correlated positively with advancing age and severity of liver disease, and certain variants may contribute to the persistence of HBV infection. ***J. Med. Virol. 68:50–59, 2002.*** © 2002 Wiley-Liss, Inc.

KEY WORDS: T cell epitope; B cell epitope; *a* determinant

INTRODUCTION

Hepatitis B virus (HBV) infection is a major health problem worldwide, affecting approximately 350 million people. The clinical consequences of HBV infection range from acute self-limited infection or fatal fulminant hepatitis, to various forms of chronic infection, including inactive carrier, chronic hepatitis, cirrhosis, and hepatocellular carcinoma (HCC) [Lee, 1997]. The humoral immune responses to HBV surface (S) protein contribute to the clearance of circulating HBV particles, whereas cellular immune responses are responsible for the elimination of infected hepatocytes [Chisari and Ferrari, 1995]. Alternatively, the emergence of variant viruses that escape from humoral or cellular immunity may be implicated in the persistence or progression of HBV infection [Wakita et al., 1991; Ehata et al., 1992; Chuang et al., 1993]. The role of immune escape variants in the natural course of chronic HBV infection remains largely unknown [Zuckerman and Zuckerman, 1999].

The production of an antibody to the *a* determinant (amino acids 124–148) located within the major hydrophilic region of the S gene after vaccination usually

Grant sponsor: Department of Health, National Science Council; Grant number: NSC 90-2314-B-002-182-MH; Grant sponsor: Executive Yuan, and National Taiwan University Hospital.

*Correspondence to: Prof. Ding-Shinn Chen, Dean, College of Medicine, National Taiwan University,1 Chang-Te St., Taipei 100, Taiwan. E-mail: dschen@ha.mc.ntu.edu.tw

Accepted 5 March 2002

DOI 10.1002/jmv.10169

Published online in Wiley InterScience (www.interscience.wiley.com)

confers protection against HBV infection [Wallace and Carman, 1997]. There is ample evidence that escape mutations within the *a* determinant occur under strong selective pressure by vaccination in infants born to hepatitis B surface antigen (HBsAg)-positive mothers or by immunoprophylaxis in liver transplantation recipients [Carman et al., 1990, 1996; Hsu et al., 1997]. S gene mutants can also be found in patients receiving lamivudine treatment, accompanying mutation of the HBV polymerase gene [Yeh et al., 2000]. In addition to the *a* determinant, the S gene of HBV contains putative HLA class I-restricted cytotoxic T lymphocyte (CTL) epitope (amino acids 28–51) [Chisari and Ferrari, 1995; Tai et al., 1997]. In general, viral mutations in CTL epitope may evade cellular immunity and thus contribute to persistent infection [Pircher et al., 1990; Wakita et al., 1991; Ehata et al., 1992; Hunt et al., 2000]. Little is known about the prevalence of escape mutations within the *a* determinant and CTL epitope in the natural course of chronic HBV infection. A recent study revealed that mutations at amino acid positions 40 and 47 of the S protein that coincides with the HLA-restricted CTL epitope occurred at a high frequency in patients with chronic hepatitis B [Tai et al., 1997]. Nevertheless, the few patients analyzed in this study were limited to those with HCC. Another study showed that mutations in the *a* determinant occurred commonly in the natural course of HBV infection, but mainly within the first loop (amino acids 124–137) of this determinant, unlike the mutations induced by immunization [Ogura et al., 1999]. The patients in the study, however, were mostly infected with serotype *adr* and the mutational patterns of the HBV S protein in patients with hepatitis B-related cirrhosis and HCC were not evaluated. These lines of evidence prompted us to investigate the prevalence and clinical significance of variations within the *a* determinant and HLA-restricted CTL epitope among different HBV serotypes in Taiwan where chronic HBV infection is hyperendemic.

In the present study, the prevalence of the naturally occurring variations was examined within the *a* determinant and the HLA-restricted CTL epitope of the S gene in patients infected with HBV serotype *adw* or *adr*. To elucidate the correlation between variations within the *a* determinant or the HLA-restricted T cell epitope and the clinical stages of chronic HBV infection, patients with different clinical outcomes were studied, including asymptomatic carrier, chronic hepatitis, circrhosis, and HCC.

MATERIALS AND METHODS

Patients

The study design was cross-sectional and observational. Serum samples from different groups of patients with various forms of chronic HBV infection were obtained from the National Taiwan University Hospital and were used for virological analysis. The four groups of patients consisted of 33 asymptomatic HBsAg carriers (Group I), 31 patients with chronic hepatitis (Group II), 22 patients with cirrhosis (Group III), and 36 patients with HCC (Group IV), and were all seropositive for HBV DNA by polymerase chain reaction (PCR). The diagnosis of different stages of liver disease was based upon clinical and biochemical data, serum α-fetoprotein level and ultrasonographic findings supplemented by histopathological examinations. None of the patients had concomitant hepatitis C virus, hepatitis D virus or human immunodeficiency virus-1 infection.

Extraction of Serum HBV DNA and Determination of S Gene Sequences

Serum viral DNA was extracted by using kits available commercially (QIAamp DNA Blood Mini Kit, QIAGEN Inc., Valencia, CA). The part of the S gene encoding both the HLA-restricted CTL epitope (amino acids 28–51, region T) and the *a* determinant (amino acids 124–148, region A) was amplified by single-round or nested PCR. The outer primers were ST1 (nucleotides [nt] 159–178: 5′-AGAACATCGCATCAGGACTC-3′) and SB2 (nt 642–623: 5′-CATAGGTATCTTGCGA-AAGC-3′), and the inner primers were ST3 (nt 181–200: 5′-AGGACCCCTGFCGTGTTAC-3′) and SB4 (nt 619–600: 5′-AGATGATGGGATGGGAATAC-3′). The amplified products were subsequently sequenced directly to determine the nucleotide sequences with an automatic sequencer (Model 377A, Applied Biosystems, Foster City, CA). To avoid false positivity of PCR, precautions described by Kwok and Higuchi [1989] were followed strictly.

Identification of HBV Serotype and Variations of S gene

The amino acid consensus sequences of the region T and the region A of HBV S gene for both serotypes *adw* and *adr*, used in a previous study [Tai et al., 1997], were adopted. The amino acid sequences deduced from the nucleotide sequences were then compared to those consensus sequences and classified into serotype *adw* or *adr*. The *adw* and *adr* consensus sequences were identical to the corresponding consensus sequence derived from our 85 *adw* and 37 *adr* amino acid sequences. The amino acid consensus sequences outside the regions A and T (amino acids 22–27 and 52–123, region O) were derived from the 85 *adw* and 37 *adr* amino acid sequences of the present study, respectively (Fig. 1A). Variations were defined as the presence of amino acids that differed from the corresponding consensus sequences (*adw* or *adr*).

Statistical Analysis

The frequencies of amino acid variation over the three regions (A, T, and O) per site per amino acid sequence were calculated and then compared among different disease stages (Group I–IV), gender, serotype (*adw* or *adr*) of HBV, and age (<50 or >50 years old). Because the frequency of amino acid variation was low for most viral sequences, the risk of variation was

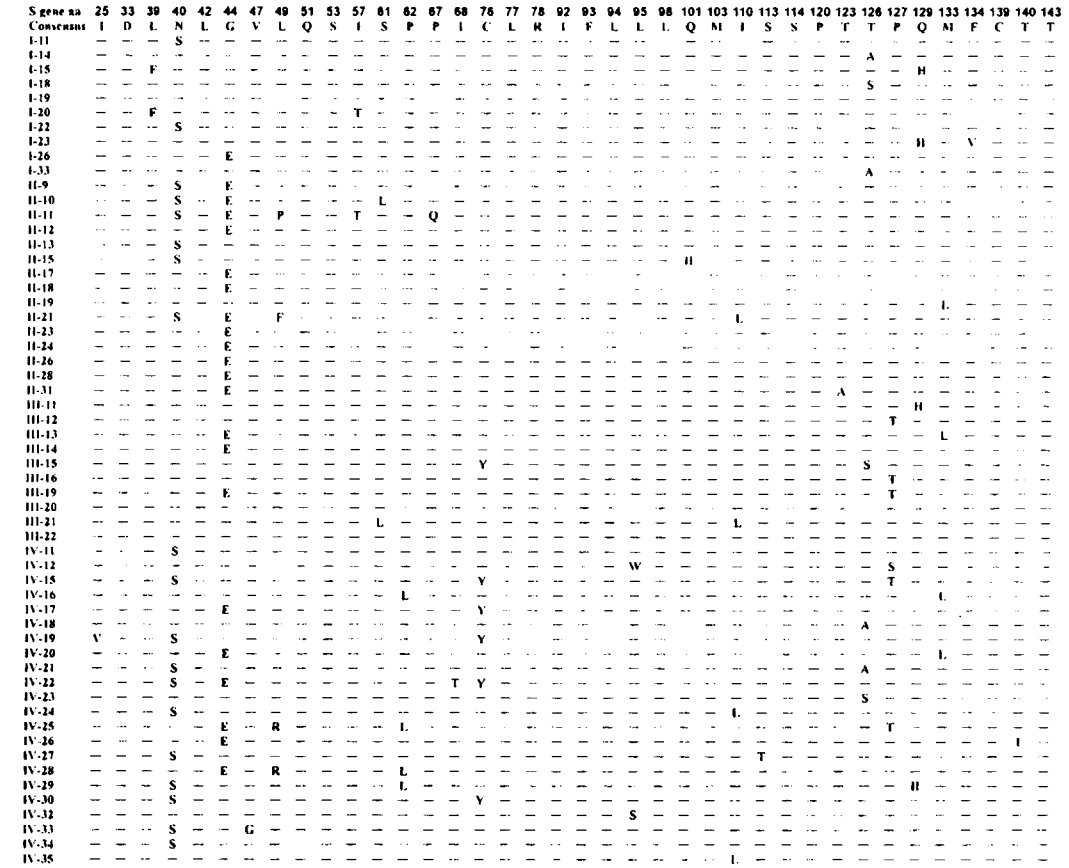
A

```
                aa 22        28                                  51
adw        LTKILTIPQS LDSWWTSLNF LGGTPVCLGQ NSQSQISSHS PTCCPPICPG YRWMCLRRFI
adr        --R------- ---------- ---A-T-P-- ----PT-N-- --S------- ----------

                                                              124
adw        IFLCILLLCL IFLLVLLDYQ GMLPVCPLIP GSSTTSTGPC KTCTTPAQGT SMFPSCCCTK
adr        ---F------ ---------- ---------- -T-------- ----I----- ----------

                    148
adw        PTDGNCT
adr        -S-----
```

B ***adw***



adr

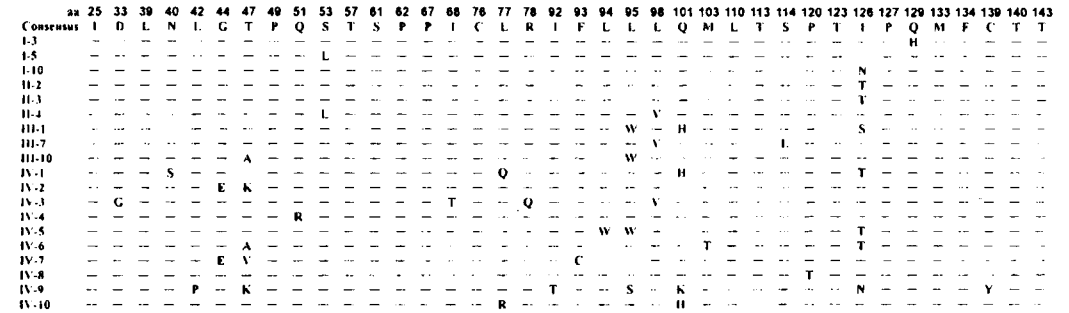


Fig. 1. **A:** Amino acid consensus sequences of partial surface protein (amino acids 22–148) of HBV serotype *adw* and *adr*, encompassing a putative HLA-restricted cytotoxic T lymphocyte (CTL) epitope (amino acids 28–51) and the *a* determinant (amino acids 124–148). Dashes indicate identity to the amino acid residues of the consensus sequence of HBV serotype *adw*. **B:** Amino acid sequence alignment of partial surface protein in 85 patients infected with HBV serotype *adw* and 37 patients infected with HBV serotype *adr* (GenBank accession numbers: AF293982–AF294006, AF463898–AF463994). HBVs whose amino acid sequence being identical to the consensus sequence, and their hosts, were not listed in this alignment. Codon numbers and amino acid residues of the corresponding *adw* or *adr* consensus sequence are listed in the top two rows. Amino acid residues are expressed in universal genetic codes. Dashes indicate identity to the corresponding amino acid consensus sequence of HBV serotype *adw* or *adr*. aa, amino acid.

compared by using the Poisson model. Because of the correlated nature (3 sites per sequence) of observed samples, generalized estimating equation (GEE) model was used to correct the statistical inference. The association between amino acid variation and variable parameters was measured by using the relative rate of variation (RR), and its 95% confidence interval. In all analyses, a two-tailed $P < 0.05$ or 95% confidence interval (CI) not covering 1.0 was considered statistically significant.

RESULTS

Characteristics of Patients

The demographic data and the distribution of HBV serotypes among different clinical groups are shown in Table I. Males predominated in all groups of patients, especially in those >50 years and those with HCC. The mean ages among Group I–III patients were below 50 years and not significantly different from each other. Of Group IV patients, the age was >50 years in all patients infected with serotype *adr* and in 67% of those infected with serotype *adw*.

All the HBV strains in the present study belonged to serotype *adw* or *adr* because the amino acid was lysine at position 122 of the major S protein in each sequence. Serotype *adw* was predominant and accounted for 70% (85/122) of all HBV strains. The distribution of HBV serotype *adw* and *adr* was not significantly different among the four groups of patients.

Amino Acid Variations in the Encoded S Region of HBV

The locations, frequencies and nature of amino acid variation of both serotypes *adw* and *adr* among Group I to IV patients are shown in Figure 1B and Figure 2. The overall frequency of amino acid variation per site per sequence (FEQ) within the studied region of HBV S gene was 1.21%. This frequency was similar between *adr* strains and *adw* strains (1.27% vs. 1.18%). In general, the FEQ increased from Group I through IV (0.55%, 1.0%, 0.91%, and 2.17%, respectively). According to the GEE with Poisson model, the FEQ correlated significantly with advancing age, the presence of HCC and the locations of putative immunogenic epitopes, but not with gender or HBV serotypes. In each group of patients, the FEQ generally increased with advancing age, irrespective of the HBV serotypes (data not shown). The relative FEQ was 1.53 times higher in patients >50 years old ($P = 0.045$, Table II). Virus strains isolated from HCC patients had a 2.1 times higher frequency of variation than that of non-cancerous patients ($P < 0.01$, Table II). Sub-regional analysis revealed that the relative FEQ within region T or A over region O was 3.68 and 2.00 times higher, respectively (both $P < 0.01$, Table II).

Interestingly, in HCC patients infected with serotype *adw*, the frequency of amino acid variation among nine patients <50 years old was higher than that of those >50 years old ($P = 0.04$, Table II), and this age effect was significantly different from that among non-HCC patients ($P = 0.04$). This finding was not observed in Group IV patients infected with serotype *adr*, however, possibly because all patients in this group were >50 years old.

Variations Within a Known T Cell Epitope (HBsAg 28–51)

The FEQ was highest within region T for serotype *adw* (region T vs. A vs. O, 2.4% vs. 1.1% vs. 0.4%, respectively). In addition, more than half of the exchanges among region T for serotype *adw* clustered at amino acid positions 40 (sN40S, 19/85 sequences; 22%) and 44 (sG44E, 22/85 sequences; 26%) and variation of these two amino acids occurred mainly among Group II–IV patients (Fig. 2B). The FEQ was also highest within region T for serotype *adr*, although statistically not significant (region T vs. A vs. O, 1.2% vs. 1.1% vs. 0.9%, respectively). In contrast to serotype *adw*, variations of amino acid at positions 40 and 44 occurred

TABLE I. Clinical Characteristics of Patients and Distribution of Hepatitis B Virus (HBV) Serotypes According to Different Liver Disease Stages

HBV serotype	Group[a]	Number	Gender (M/F)	Mean age (years)	FEQ[b] (%)
adw	I	23	14/9	37	0.57
	II	24	18/6	39	1.13
	III	12	9/3	40	1.00
	IV	26	21/5	56	1.85
	Total	85	62/23	44	1.18
adr	I	10	5/5	37	0.50
	II	7	4/3	32	0.57
	III	10	7/3	46	0.80
	IV	10	10/0	59	3.00
	Total	37	26/11	45	1.27

[a]Groups: I, asymptomatic HBV carriers; II, chronic hepatitis; III, liver cirrhosis; IV, hepatocellular carcinoma.
[b]FEQ, frequency of amino acid variation per site per amino acid sequence.

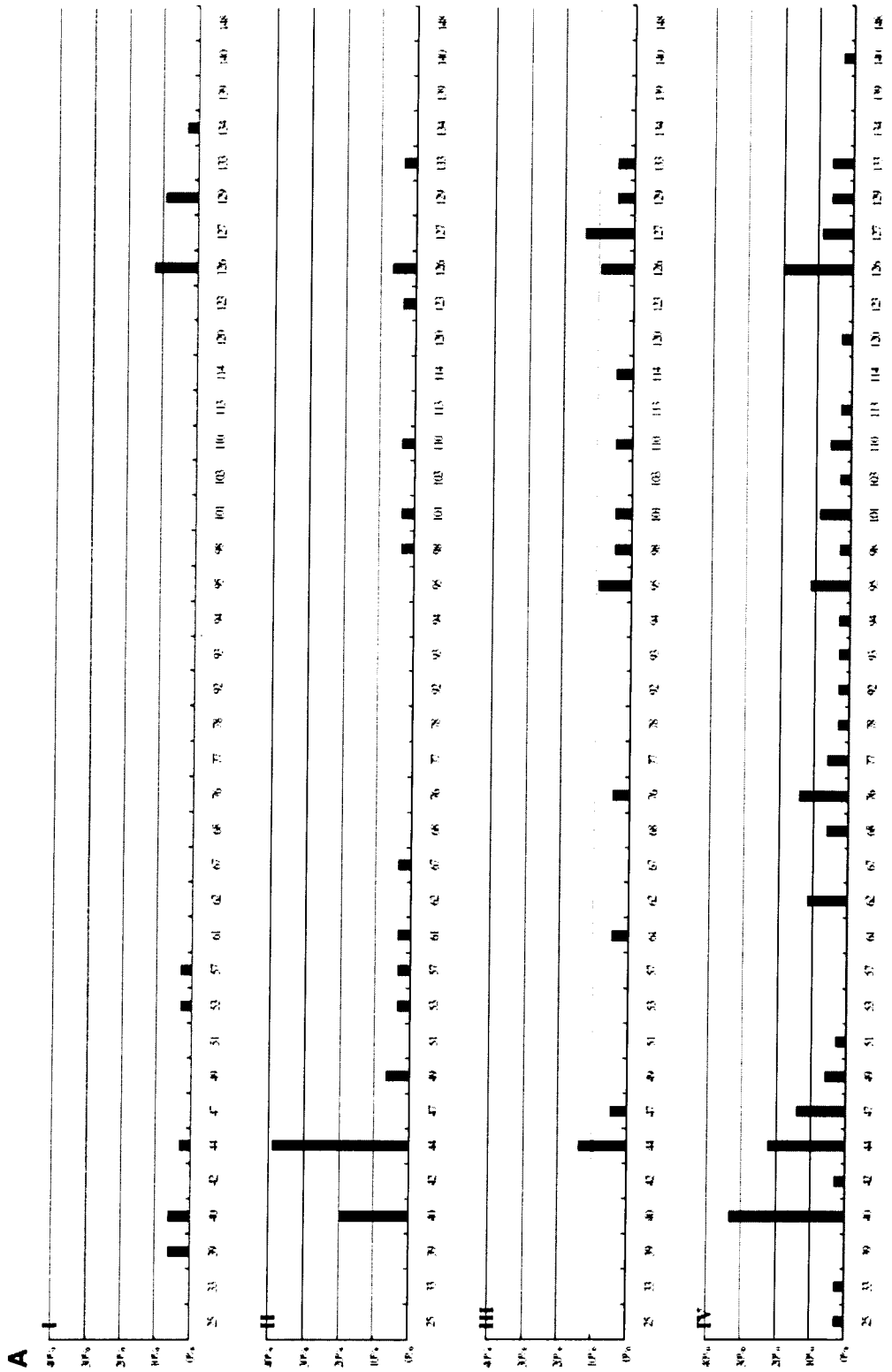


Fig. 2. Locations and frequencies of amino acid variations along all encoded surface protein of hepatitis B virus among patients with asymptomatic infection (Group I), chronic hepatitis (Group II), cirrhosis (Group III) and hepatocellular carcinoma (Group IV). **A:** All patients ($n = 122$). **B:** Patients infected with serotype *adw* ($n = 85$). **C:** Patients infected with serotype *adr* ($n = 37$). Variations are defined as the presence of amino acid different from the *adw* or *adr* amino acid consensus sequence. Horizontal axis indicates the amino acid position of encoded major surface protein (amino acids 22–148). Vertical axis indicates the frequency of amino acid variation at each amino acid position. Only those positions with amino acid variation are shown. The *a* determinant of major surface protein coincides within amino acids 124–148. A putative HLA class I-restricted cytotoxic T lymphocyte epitope coincides within amino acids 28–51.

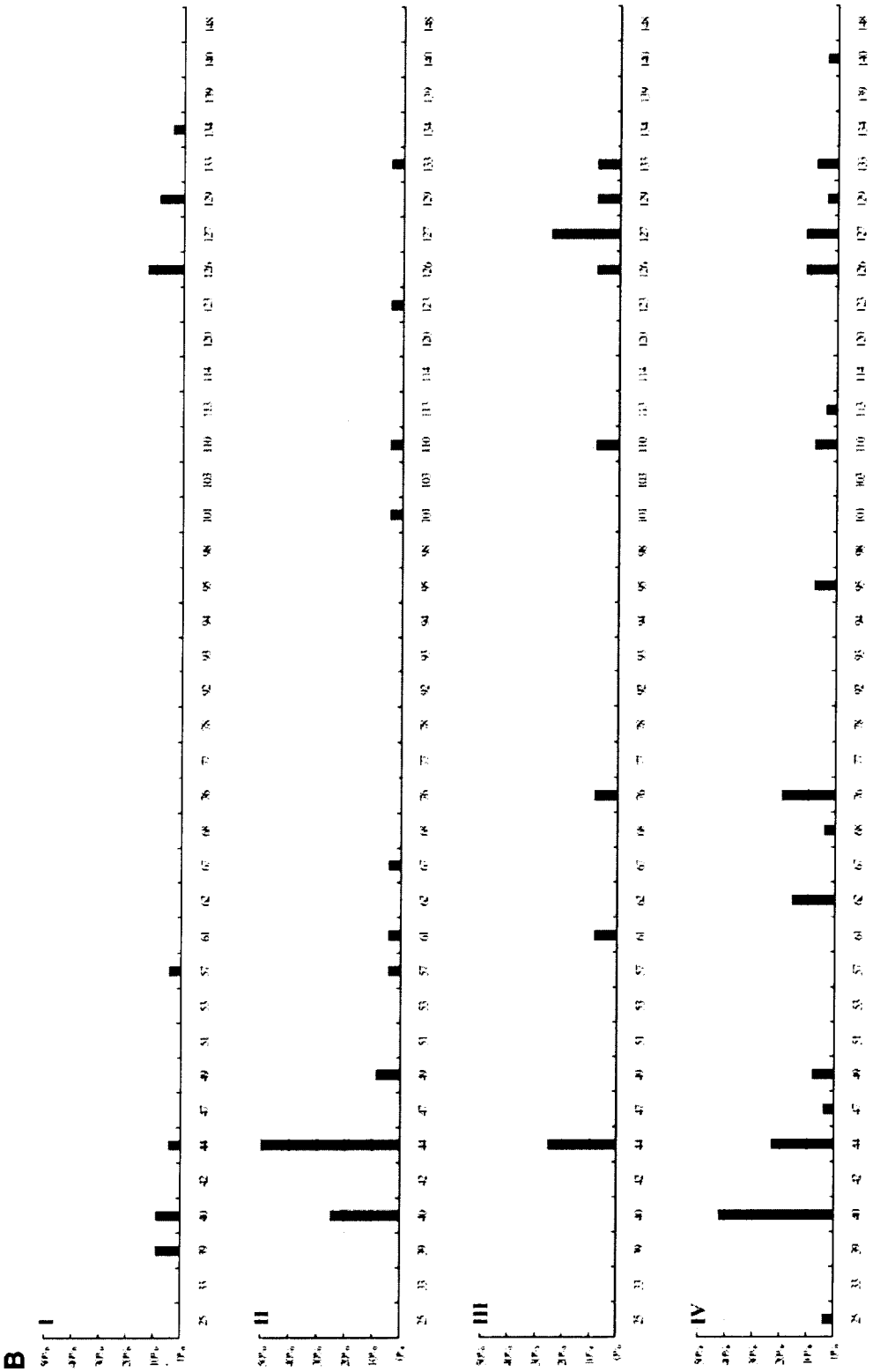


Fig. 2. (*Continued*)

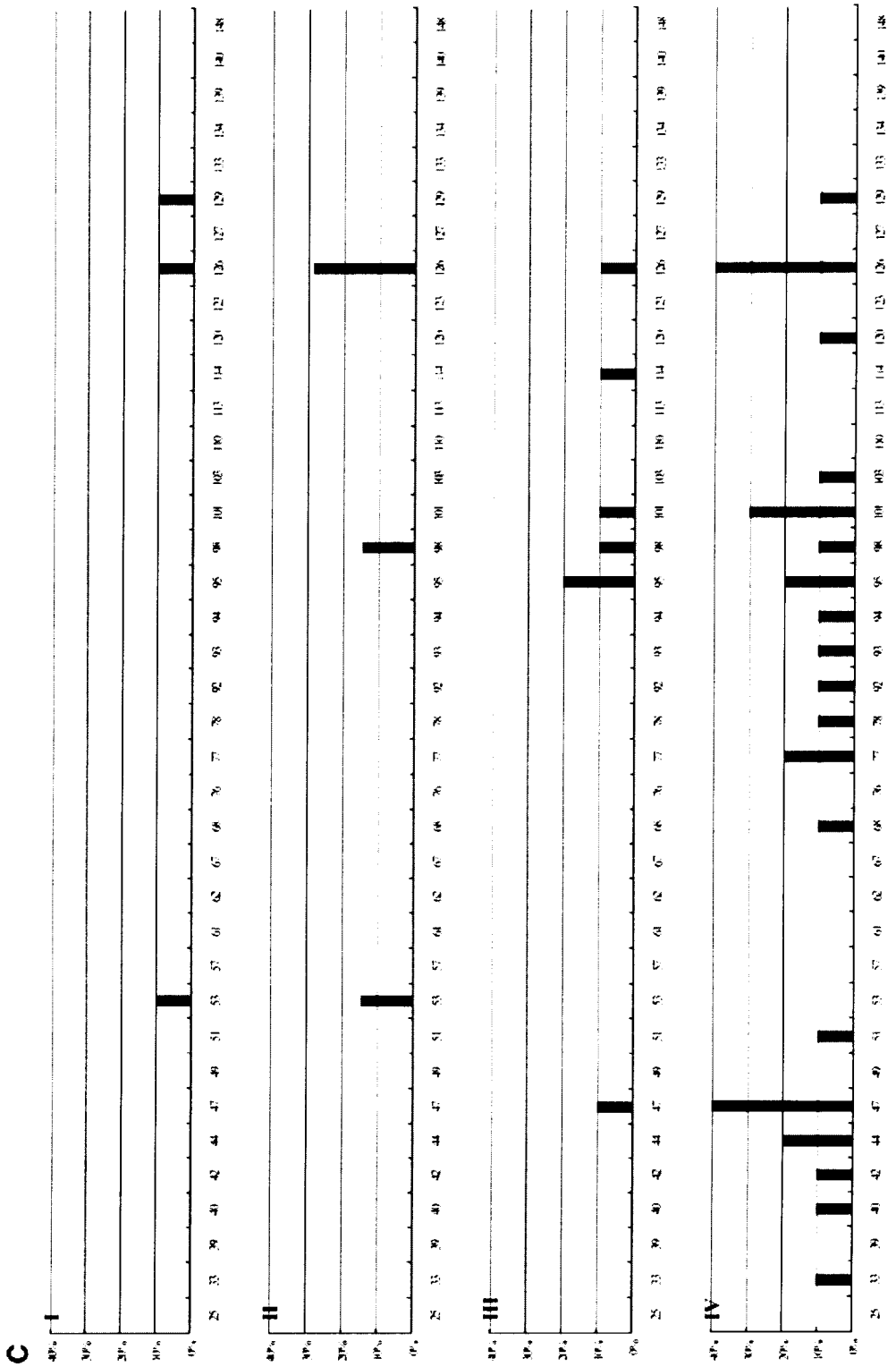


Fig. 2. (*Continued*)

TABLE II. Relative Frequency of Amino Acid Variation of Hepatitis B Virus (HBV) Surface Protein for all Patients and Those Infected With Serotype *adw*

	Parameters[a]	Relative frequency	95% CI[b]	*P*-value
All patients	Region T vs. O	3.68	2.63–5.14	0.000
	Region A vs. O	2.00	1.34–1.99	0.001
	HCC[c] vs. non-HCC	2.10	1.39–3.18	0.000
	Age > 50 vs. ≤50	1.53	1.01–2.33	0.045
Patients infected with HBV serotype *adw*	Region T vs. O	5.90	4.12–8.45	0.000
	Region A vs. O	2.77	1.62–4.74	0.000
	HCC[c] vs. non-HCC	2.50	1.36–4.60	0.003
	Age > 50 vs. ≤50	2.22	1.29–3.84	0.004
	HCC age > 50 vs. ≤50[d]	0.44	0.20–0.96	0.040

[a]All parameters correlated with the frequency of amino acid variation are analyzed by the generalized estimating equation with Poisson model.
[b]95% CI means 95% confidence interval.
[c]HCC, hepatocellular carcinoma.
[d]Indicates the interaction effect of HCC and age parameters on the relative frequency of amino acid variation. For HCC patients infected with serotype *adw*, the relative frequency of amino acid variation is inversely correlated with age >50 years.

uncommonly for serotype *adr*. Instead, amino acid 47 varied commonly (40%) within region T in HCC patients infected with serotype *adr* (Fig. 2B).

Variations Within a Known B Cell Epitope (HBsAg 124–148)

The amino acids within the encoded second loop (HBsAg 138–148) of the *a* determinant of the S gene were highly conserved for both serotypes *adw* and *adr*. Most variations of amino acid clustered within the first loop of the *a* determinant, especially at amino acid position 126. In addition, the amino acid variations within this region were also encountered commonly in asymptomatic carriers. The variation at amino acid 145 of S protein, an important target at the second loop (amino acids 139–147) of the *a* determinant that is responsible for most hepatitis B immunoprophylaxis failure [Hsu et al., 1997; Wallace and Carman, 1997], was not observed in this study.

DISCUSSION

In this cross-sectional study, by correlating amino acid variations within the *a* determinant and the HLA-restricted T cell epitope with different disease stages of chronic HBV infection, the overall frequency of variation per site per amino acid sequence was found to be 1.21%. The frequency was found to correlate positively with the advancing age and the severity of liver disease. This trend of amino acid variation held true for both serotypes *adw* and *adr*. The variation rate of HCC patients aged ≤50 years who were infected with serotype *adw* was significantly higher than that of age-matched non-HCC patients. In addition, it was found that the encoded second loop of the *a* determinant was highly conserved and the humoral escape variation at position 145 of major S protein did not occur in the natural course of the Taiwanese hepatitis B carriers. It was also observed that hotspot mutations differed between serotype *adw* (amino acid positions 40 and 44 of the major S protein) and *adr* (position 47) during the progression of chronic hepatitis B.

HBV variants could display alteration of epitopes important in the host immune recognition, enhanced virulence with increased levels of HBV replication, resistance to antiviral treatments or facilitated cell attachment/penetration [Lok et al., 1994; Chisari and Ferrari, 1995; Buckwold et al., 1996; Hunt et al., 2000]. In the natural course of HBV infection, cellular and humoral immune pressure against virus-specific proteins in an attempt to clear the virus may select fittest viral strains with survival advantages. Several studies have shown that patients with chronic active hepatitis often carry missense mutations or deletions in the middle of the HBV core gene, whereas similar mutations are not observed among patients with inactive disease, suggesting that these mutations in the core region may be CTL-specific escape mutations [Wakita et al., 1991; Ehata et al., 1992; Chuang et al., 1993]. In addition to the core protein, prior studies also found a high prevalence of variations within the putative HLA class I-restricted CTL epitope and the *a* determinant of HBV S protein [Chisari and Ferrari, 1995; Tai et al., 1997; Ogura et al., 1999; Chen and Oon, 1999a,b]. Our results showed consistently that the frequency of amino acid variation was highest within this putative T cell epitope. In addition, variations at amino acid positions 40 and 44 occurred more frequently among patients with chronic hepatitis (Groups II–IV) but rarely among asymptomatic carriers (Group I). It is therefore suggested that the hotspots 40 and 44 variants may have been selected by CTL-mediated host immune response in the early course of serotype *adw* infection and may play a role in the persistence of HBV infection. Amino acid 47 variation was observed in five of the 37 (13.5%) serotype *adr* but in only two of the 85 (2.5%) serotype *adw*. Of interest, four of the five variations at position

47 occurred in Group IV patients with serotype *adr* infection. Therefore, the variation at position 47 might be a late event in the natural course of serotype *adr* infection, which could possibly explain the low prevalence of amino acid substitution at position 47 in a previous study that most of their subjects were limited to asymptomatic carriers or patients with chronic hepatitis [Ogura et al., 1999]. Taken together, these findings suggest that this region is likely an immunogenic epitope. Furthermore, there might be a hierarchical priority for amino acid to vary and these variations may be serotype-specific and disease stage-correlated.

Amino acid 145 within the second loop of *a* determinant has attracted much attention, because substitution of the amino acid at this position may impose a negative impact on the efficacy of hepatitis B vaccination or immunoprophylaxis with hepatitis B immunoglobulin (HBIG) for liver transplantation [Zuckerman, 2000]. Several studies showed that glycine-to-arginine substitution at this position was associated strongly with a breakthrough of the HBV infection after hepatitis B vaccination or HBIG immunoprophylaxis [Carman et al., 1990, 1996]. Whether amino acid 145 variant could emerge in the natural evolution of long-term HBV infection in hepatitis B endemic areas should be clarified and this information will help in designing further preventive strategies against HBV infection [Hsu et al., 1999; Ogata et al., 1999; Zuckerman, 2000]. In this study, we did not find any amino acid 145 variant in our patients; however, the methods used and results cannot exclude the possibility that amino acid 145 variant exists as a minor population, which may not be detected by PCR and direct sequencing.

The results indicate that the frequency of amino acid change per site per sequence was higher for patients >50 years old, irrespective of the clinical groups of patients or HBV serotypes. Because most HBsAg carriers in Taiwan acquire their HBV infection in the perinatal period or infancy, the duration of HBV infection is assumed to approximate their age. HBV variants are believed to arise by random mutation followed by host immune selection, and thus the increased prevalence of amino acid change within the encoded S protein for older subjects can possibly be accounted for by the long-term HBV evolution. Nevertheless, the substitution rate was strikingly high among young HCC patients infected with serotype *adw*, and even greater than that among old HCC patients with serotype *adw* infection. Accordingly, certain HBV serotype *adw* strains, whether possessing these amino acid changes from the beginning of HBV infection or developing variations soon during the course of infection may exist, and might be associated with the early development of HCC. These facts also support the previous observation that young Taiwanese HCC patients are infected mainly with HBV genotype B (all are serotype *adw*) [Kao et al., 2000]. Further biological studies of these HBV strains, particularly on the integrated HBV-DNA within both the HCC and non-HCC liver tissue are needed to clarify the mechanisms of this association.

For both serotypes *adw* and *adr*, amino acid variation occurred frequently within the region outside the CTL epitope and the *a* determinant. This region did not overlap with the major catalytic domains of HBV polymerase gene. Therefore, random mutations within this region might occur without affecting the ability of the virus to survive. Some variations, however, clustered between positions 77–103 for serotype *adr*. Previous studies also revealed that amino acid variations could be occasionally found within this region [Ogura et al., 1999; Weinberger et al., 2000]. Although this region is a well characterized transmembrane region by structural analysis, in vitro immunogenic studies suggested that peptide spanning amino acids 88–106 of major surface protein might be HLA-A2-restricted epitopes [Chisari and Ferrari, 1995]. Because the consensus sequences were determined upon a limited sequence data, the HLA profiles of these patients were not determined, and long-term evolution of this S gene fragment was not clarified, whether these amino acid substitutions represent HLA-restricted or HBV subtype-specific variations, or are truly immune escape variations remains to be explored.

In summary, the results suggest that certain immune escape variations may contribute to the persistence of HBV infection. The mutation hotspots of the major S protein differ between serotypes *adw* and *adr*, and may emerge during different disease stages of chronic hepatitis B. The time of emergence and the role of these immune escape variants in the pathogenesis of chronic hepatitis B await long-term longitudinal studies. Finally, further biological and genomic studies are needed to clarify whether particular HBV strains are associated with the development of HCC in young patients.

ACKNOWLEDGMENTS

The study was supported by grants from the Department of Health, the National Science Council (NSC 90-2314-B-002-182-MH), Executive Yuan, and National Taiwan University Hospital, Taiwan. C.J.L. is supported by a research scholarship from the Academia Sinica, Taiwan.

REFERENCES

Buckwold VE, Xu Z, Chen M, Yen TSB, Ou JH. 1996. Effects of a naturally occurring mutation in the hepatitis B virus basal core promoter on precore gene expression and viral replication. J Virol 70:5845–5851.

Carman WF, Zanetti AR, Karayiannis P, Waters J, Manzillo G, Tanzi E, Zuckerman AJ, Thomas HC. 1990. Vaccine-induced escape mutant of hepatitis B virus. Lancet 336:325–329.

Carman WF, Trautwein C, van Duersen FJ, Colman K, Dorman E, McIntyre G, Waters J, Kliem V, Muller R, Thomas HC, Manns MP. 1996. Hepatitis B virus envelope variation after transplantation with and without hepatitis B immune globulin prophylaxis. Hepatology 24:489–493.

Chen WN, Oon CJ. 1999a. Mutation "hot spot" in HLA class I-restricted T-cell epitope on hepatitis B surface antigen in chronic carriers and hepatocellular carcinoma. Biochem Biophys Res Commun 262:757–761.

Chen WN, Oon CJ. 1999b. Human hepatitis B virus mutants: significance of molecular changes. FEBS Lett 453:237–242.

Chisari FV, Ferrari C. 1995. Hepatitis B immunopathogenesis. Ann Rev Immunol 13:29–60.

Chuang WL, Omata M, Ehata T, Yokosuka O, Ito Y, Imazeki F, Lu SN, Chang WY, Ohto M. 1993. Precore mutations and core clustering mutations in chronic hepatitis B virus infection. Gastroenterology 104:263–271.

Ehata T, Omata M, Yokosuka O, Hosoda K, Ohto M. 1992. Variations in codons 84–101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. J Clin Invest 89:332–338.

Hsu HY, Chang MH, Ni YH, Lin HH, Wang SM, Chen DS. 1997. Surface gene mutants of hepatitis B virus in infants who develop acute or chronic infections despite immunoprophylaxis. Hepatology 26:786–791.

Hsu HY, Chang MH, Liaw SH, Ni YH, Chen HL. 1999. Changes of hepatitis B surface antigen variants in carrier children before and after universal vaccination in Taiwan. Hepatology 30:1312–1317.

Hunt CM, McGill JM, Allen MI, Condreay LD. 2000. Clinical relevance of hepatitis B viral mutations. Hepatology 31:1037–1044.

Kao JH, Chen PJ, Lai MY, Chen DS. 2000. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. Gastroenterology 118:554–559.

Kwok S, Higuchi R. 1989. Avoid false positive with PCR. Nature 339:237–238.

Lee WM. 1997. Hepatitis B virus infection. N Engl J Med 337:1733–1745.

Lok ASF, Akarca U, Greene S. 1994. Mutations in the pre-core region of hepatitis B virus serve to enhance the stability of the secondary structure of the pre-genome encapsidation signal. Proc Natl Acad Sci USA 91:4077–4081.

Ogata N, Cote PJ, Zanetti AR, Miller RH, Shapiro M, Gerin J, Purcell RH. 1999. Licensed recombinant hepatitis B vaccines protect against infection with the prototype surface gene mutant of hepatitis B virus. Hepatology 30:779–786.

Ogura Y, Kurosaki M, Asahina Y, Enomoto N, Marumo F, Sato C. 1999. Prevalence and significance of naturally occurring mutations in the surface and polymerase genes of hepatitis B virus. J Infect Dis 180:1444–1451.

Pircher H, Moskophidis D, Rohrer U, Burki K, Hengartner H, Zinkernagel RM. 1990. Viral escape by selection of cytotoxic T cell-resistant virus variants in vivo. Nature 346:629–633.

Tai PC, Banik D, Lin GI, Pai S, Pai K, Lin MH, Yuoh G, Che S, Hsu SM, Chen TC, Kuo TT, Lee CS, Yang CS, Shih C. 1997. Novel and frequent mutations of hepatitis B virus coincide with a major histocompatibility complex class I-restricted T-cell epitope of the surface antigen. J Virol 71:4852–4856.

Wakita T, Kakumu S, Shibata M, Yoshioka K, Ito Y, Shinagawa T, Ishikawa T, Takayanagi M, Morishima T. 1991. Detection of pre-C and core region mutants of hepatitis B virus in chronic hepatitis B virus carriers. J Clin Invest 88:1793–1801.

Wallace LA, Carman WF. 1997. Surface gene variations of HBV: scientific and medical relevance. J Viral Hepat 3:5–16.

Weinberger KM, Bauer T, Böhm S, Jilg W. 2000. High genetic variability of the group-specific a-determinant of hepatitis B virus surface antigen (HBsAg) and the corresponding fragment of the viral polymerase in chronic virus carriers lacking detectable HBsAg in serum. J Gen Virol 81:1165–1174.

Yeh CT, Chien RN, Chu CM, Liaw YF. 2000. Clearance of the original hepatitis B virus YMDD-motif mutants with emergence of distinct lamivudine-resistant mutants during prolonged lamivudine therapy. Hepatology 31:1318–1326.

Zuckerman AJ. 2000. Effect of hepatitis B virus mutants on efficacy of vaccination. Lancet 355:1382–1384.

Zuckerman AJ, Zuckerman JN. 1999. Molecular epidemiology and hepatitis B virus mutants. J Med Virol 58:193–195.