from Vietnam, Thailand, Myanmar, China, Hong Kong and HBV/Ce from Japan, Korea. The subtypes (subgenotypes) of the 16 strains were confirmed by a phylogenetic analysis of the complete genome (Fig. 2a). Taking advantage of the five SNPs of T1041, C1044, A1050, A1053 and C1155, a RFLP method with three endonucleases was developed for distinguishing between HBV/Cs and HBV/Ce. PCR products of 309 bp in size (nt 964–1272), amplified on HBV/Cs strains, were split by *Ase*I digestion into two fragments of 88 and 221 bp and/or *BstE*II digestion into two fragments of 76 and 233 bp (Fig. 3), while those on HBV/Ce strains were not. In contrast, the products of 309 bp, amplified on HBV/Ce strains, were broken down by *Nci*I digestion into two fragments of 192 and 117 bp, while those on HBV/Cs strains were not.

Total 49 HBV/C samples, consisting of 24 in Hong Kong and 25 in Japan, were examined for the specificity of the novel PCR-RFLP method. Based on the PCR-RFLP, the 24 strains from Hong Kong were classified into HBV/Cs, and the 25 from Japan were HBV/Ce. To confirm the reliability of the PCR-RFLP method, the precore region plus core gene was sequenced directly on all 49 samples. All the 24 HBV/Cs and 25 HBV/Ce samples determined by PCR-RFLP were

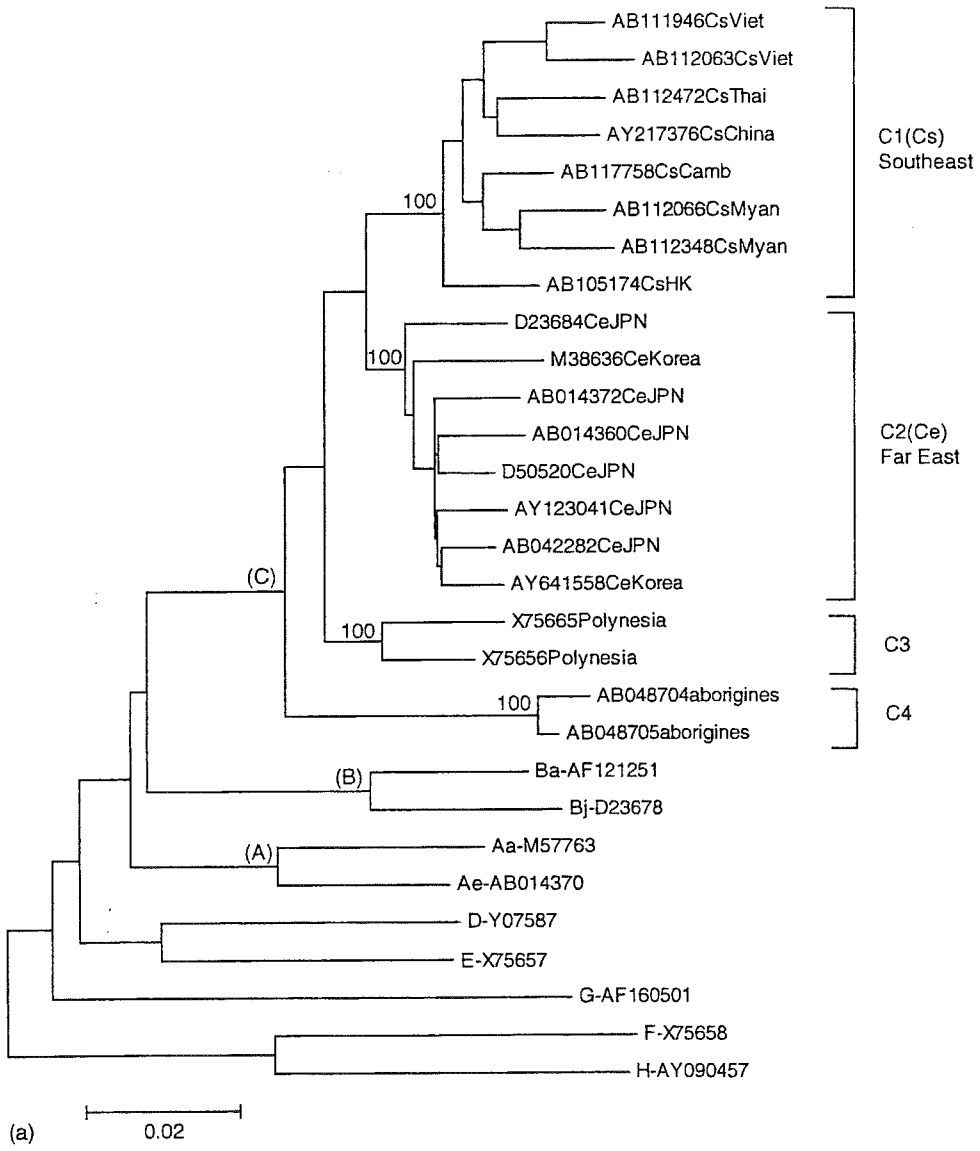

Fig. 2. (a) A phylogenetic tree constructed on the complete genome sequences of 29 HBV strains. Eight HBV/C1 (Cs) and eight HBV/C2 (Ce) strains (shown in Fig. 1) are compared along with four other HBV/C (C3 and C4) and nine HBV strains representative of the other seven genotypes (Aa, Ae, Ba, Bj, D–H). (b) A phylogenetic tree constructed on the X gene, precore and core gene sequences spanning 398 bp. Together with the above 29 representative sequences retrieved from database, 24 HBV/C1 (Cs) strains determined by PCR-RFLP belong to HBV/C1 (Cs) and 25 HBV/C2 (Ce) strains by PCR-RFLP had a cluster with the representative HBV/C2 (Ce) strains from database. All strains in this study are shown in bold. Each representative strain from the database are identified with accession numbers, followed by subtype and the country of origin in abbreviation for Cambodia (Camb), Hong Kong (HK), Japan (JPN), Myanmar (Myan) and Vietnam (Viet). The length of the horizontal bar indicates the number of nucleotide substitution per site.

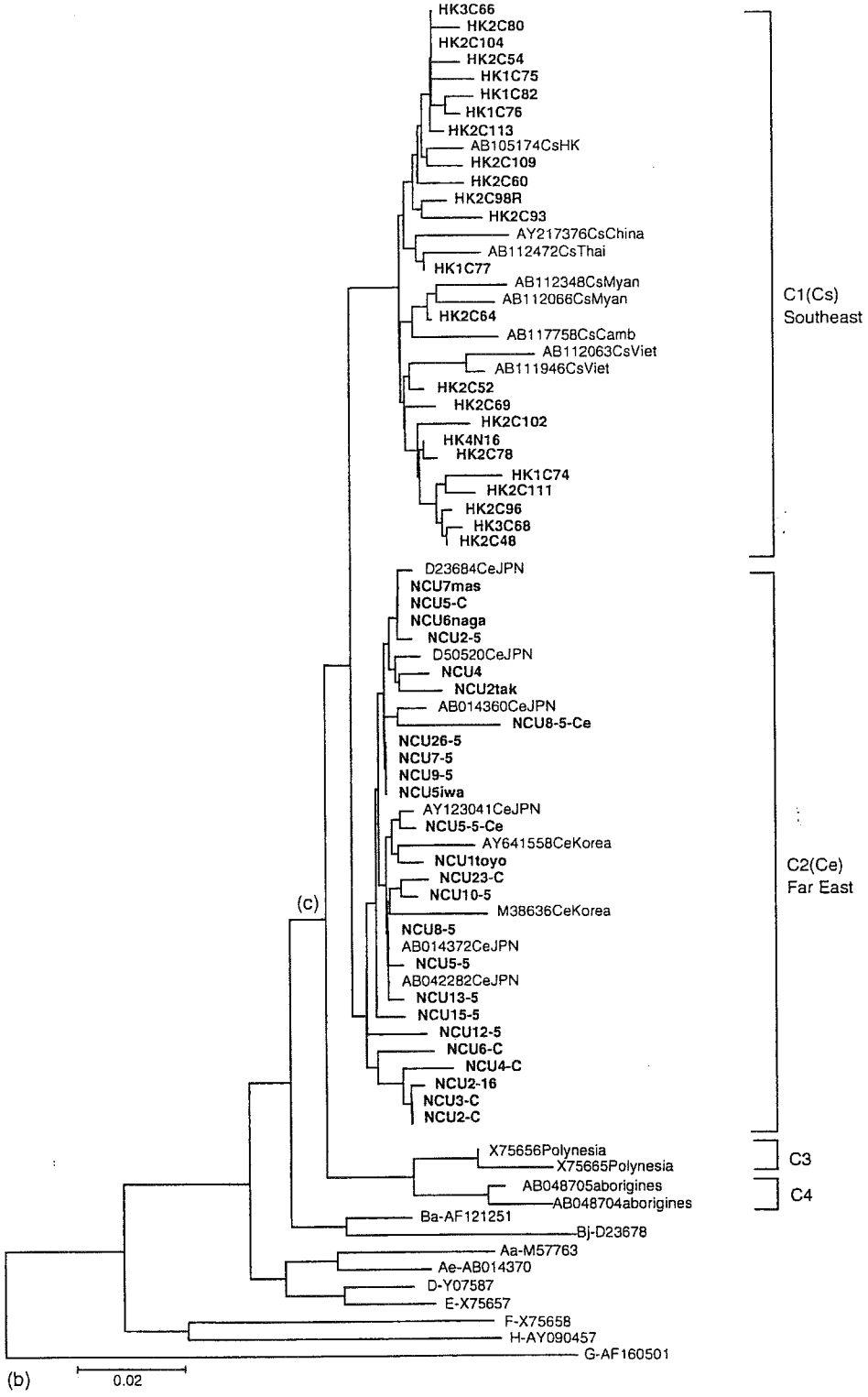

Fig. 2. (*Continued*).

completely classified into each subtype (subgenotype) by sequencing. To evaluate the sensitivity of the method, serial dilution of each HBV/Cs and HBV/Ce clones was used for the hemi-nested PCR, and its detection limit was five copies per assay.

3.3. Mutations in the enhancer, BCP and precore region in patients infected with HBV/C1 and C2

An alignment of sequences covering the BCP and the encapsidation signal (ε) in HBV/Cs and HBV/Ce allowed

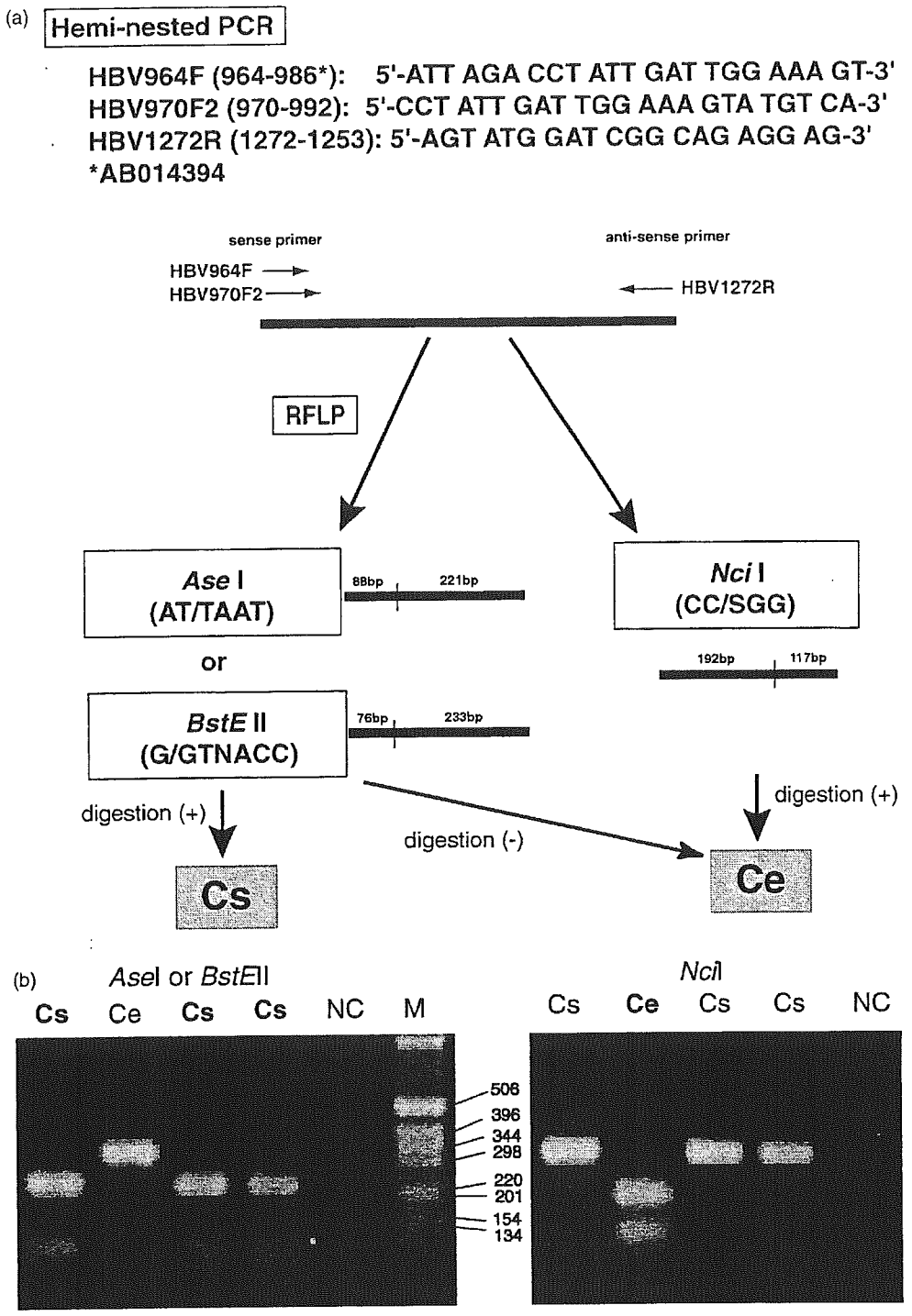

Fig. 3. (a) The strategy of a novel subtyping assay of HBV/C based on PCR-RFLP with *BstE*II, *Ase*I, *Nci*I restriction enzymes. HBV/Cs is digested by *BstE*II and/or *Ase*I, while HBV/Ce is digested by only *Nci*I. (b) Identification of restriction patterns obtained by restriction endonuclease digestion. Using hemi-nested PCR followed by cleavage with three kinds of restrict enzyme, it was possible to distinguish between HBV/Cs and HBV/Ce.

the identification of specific substitutions for HBV/C1 and HBV/C2 strains at nt 1721, 1757, 1775, 1856 and 1858 (Table 2). The prevalence of T1653, A1896 and A1899 substitutions was significantly higher in HBV/Ce than that in HBV/Cs, while the prevalence of A1727 and A1898 substitutions was higher in HBV/Cs. Double mutation in BCP (T1762/A1764) was highly prevalent in both subtypes (subgenotypes). Interestingly, the precore stop mutation (A1896), accompanied by a C-to-T substitution at nt 1858 forming a base pair with it, was found only in HBV/Ce strains (45/162, 28%), whereas no mutation was found in HBV/Cs strains due to C1858. Another precore mutation (A1898), accompanied by a C-to-T mutation at nt 1856, was found in HBV/Cs strains (7/58, 12%) (Table 2).

Table 2
Subtype-specific mutations in basic core promoter and encapsidational signal of HBV/Cs and Ce strains

Nucleotide position	Cs	This study ($n = 24$)	Database ($n = 34$)	Ce	This study ($n = 25$)	Database ($n = 137$)	P-value
1653	T	2 (8%)	0	T	7 (28%)	37 (27%)	<.0001
1721	A	22 (92%)	33 (97%)	G	24 (96%)	137 (100%)	<.0001
1727	A	19 (79%)	30 (88%)	A	13 (52%)	59 (43%)	<.0001
1757	A	13 (54%)	10 (29%)	G	25 (100%)	137 (100%)	<.0001
1762/1764	T/A	21 (88%)	13 (38%)	T/A	20 (80%)	65 (47%)	NS
1775	G	15 (63%)	28 (82%)	A	25 (100%)	132 (96%)	<.0001
1856	T	9 (38%)	6 (18%)	C	25 (100%)	137 (100%)	<.0001
1858	C	23 (96%)	23 (68%)	T	25 (100%)	137 (100%)	<.0001
1896	A	0	0	A	6 (24%)	39 (28%)	<.0001
1898	A	7 (29%)	0	A	0	0	<.0001
1899	A	1 (4%)	1 (3%)	A	4 (16%)	19 (14%)	0.029

4. Discussion

Chronic patients infected with HBV/C have a more aggressive clinical course than those infected with HBV/B [3,18]. In this study, we focused on HBV/C because it is prevalent mainly in Asia and seems to contribute to progressive liver disease and poor clinical outcomes in infected patients. Phylogenetic analyses of the complete genome show at least 4 subtypes (subgenotypes) of HBV/C (C1–4) with different geographic distribution (Fig. 2a) [19,20]. HBV/C1 was found only in Southeast Asia including Vietnam, Myanmar, Thailand, Laos, Bangladesh, Hong Kong and southern China, while HBV/C2 was found in far East Asia including Japan, Korea and northern China. Additionally, two another subtypes (subgenotypes) of HBV/C were named as C3 and C4 [19,20]. C3 was found in a large area of the Pacific from New Zealand to Polynesia, while C4 was isolated from Aborigines in Northeast Australia [17]. However, as C3 and C4 strains were rarely found in most Asian countries, we focused the classification between Cs (C1) and Ce (C2) in the present study.

A total of 118 complete genome sequences of the HBV/C strains isolated in the different geographic regions were analyzed phylogenetically in the recent study [13]; the phylogenetic subclusters within HBV/C were subsequently designated respectively to the geographic regions, i.e. "Cs" for Southeast Asian (Vietnam, Thailand, Myanmar and Southern China), and "Ce" for far East Asia (Korea, Japan, and Northern China). According to this classification, 80% of the patients in Hong Kong were belonged to the Cs and 20% to the Ce [13]. When taken in account both facts, i.e. evident geographic origins of these subtypes (subgenotypes) and the phylogenetic confirmation, the designation using the small letters (indicating possible origins) appears to be logical, similarly to the previously reported Asian "Ba" and Japanese "Bj" [5,7], Africa/Asian Aa and European Ae [8]. Hence, "Cs" and "Ce" designation was applied to the present study.

Based on five SNPs between HBV/Cs and HBV/Ce, we developed a novel PCR-RFLP method for distinguishing between HBV/Cs and HBV/Ce with high reliability. All 49 samples examined were completely classified by the PCR-RFLP. This method allows the classification between these subtypes (subgenotypes) without using expensive, labor- and time-consuming methods such as sequencing and molecular evolutionary analyses. Examining additional 171 complete sequences from database, only 9 sequences of HBV/Ce have exceptional mutations at the restriction site of *Nci*I, indicating that less than 5% of the strains known up to date are unclassified by this method, and require sequencing as previously described [17].

Some specific mutations were detected in the encapsidation signal site; the precore stop mutation (A1896), accompanied by a C-to-T substitution at nt 1858 forming a base pair with it, was found only in HBV/Ce strains, and another precore mutation (A1898), accompanied by a C-to-T mutation at nt 1856, was found only in HBV/Cs strains (Fig. 4). These mutations could stabilize the ε loop structure and the former HBeAg-negative mutants bearing a TAG stop codon mutation at codon 28 (A1896) uniformly replicate at least 20-fold better than mutants bearing a TGA stop codon at the same amino acid position enhance viral replication [21]. This C1858 variant was frequntly found in HBV/A and HBV/F [22]. Additionally, A1899 mutation was more prevalent in the HBV/Ce. As previously reported, the effects caused by these two closely linked mutations (A1896 and A1899) on viral replication are not independent each other [21]. The stringent selection for a highly efficient RNA encapsidation element may play a crucial role in the natural occurrence of these two closely linked precore mutations. Our replication model also shows that the combined mutations can induce higher replication in vitro (unpublished data). Hence, these several virological differences between the two subtypes (subgenotypes) might influence clinical outcomes such as fulminant hepatitis or hepatocarcinogemesis.

The biologic function of HBeAg remains controversial. Although HBeAg is not required for viral replication, it appears to be necessary for the establishment of chronic infection in animal models [23]. The most common mutation in the precore sequence that abrogates the synthesis of HBeAg is a stop-codon mutation (G1896A). As all HBV/Ce strains possessed T1858 and most HBV/Cs had C1858, the HBV/Cs with C1858 might be responsible for a delayed seroconversion for the loss of HBeAg in the carriers of HBV/Cs. The clinical significance of C1858 and T1858 among HBV/C

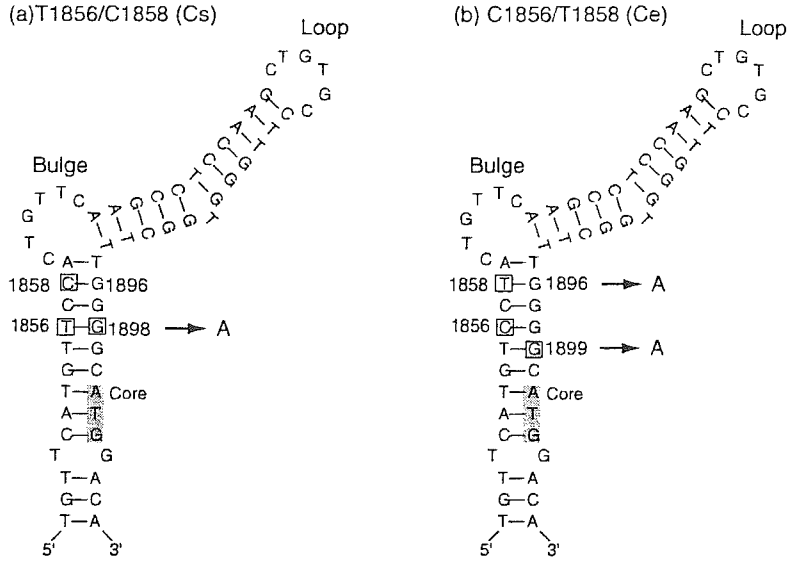

Fig. 4. Conformation of the pregenome encapsidation (ε) signal for (a) HBV/Cs and (b) HBV/Ce. The precore mutation, G1898A accompanied by T1856 forming a base pair with it, was found only in HBV/Cs strains. In contrast, the precore stop mutation, G1896A accompanied by T1858, was found only in HBV/Ce strains. A1899 mutation is significantly predominant in HBV/Ce strains.

is not well known. A previous study among multi-ethnic carriers in Hawaii indicated no significant difference in clinical characteristics between C1858 and T1858 variants [24]. However, as the number of patients was not enough to clarify the significance of this variation, further clinical studies would be required on a case–control study with large-scale cohorts.

A previous study [12] indicated that the amino acid changes specific to HBV/Cs and HBV/Ce were concentrated in the pre-S1, S and P regions, but not in the X and core regions. The pre-S1 region contains the HBV receptor for entering hepatocytes [25] and also has sites for transcriptional factors [26]. Another study [13] showed three amino acids differences in polymerase region. Therefore, the relationship between HBV/Cs and HBV/Ce and their virulence in chronic liver diseases including hepatocellular carcinoma are of great interest, since the prevalence of HBV-related hepatocellular carcinoma is extremely high in Asia compared with other regions.

In conclusion, a new PCR-RFLP method involving 5 SNPs was developed for specifically distinguishing between HBV/Cs and HBV/Ce. The two subtypes (subgenotypes) have distinct geographic distribution and virological characteristics. The novel PCR-RFLP would be useful in evaluating clinical, epidemiological and virological differences between HBV/Cs and HBV/Ce infections in countries where HBV genotype C endemic.

References

[1] Kidd-Ljunggren K, Miyakawa Y, Kidd AH. Genetic variability in hepatitis B viruses. J Gen Virol 2002;83:1267–80.

[2] Kidd-Ljunggren K, Myhre E, Blackberg J. Clinical and serological variation between patients infected with different Hepatitis B virus genotypes. J Clin Microbiol 2004;42:5837–41.

[3] 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.

[4] Chan HL, Hui AY, Wong ML, et al. Genotype C hepatitis B virus infection is associated with an increased risk of hepatocellular carcinoma. Gut 2004;53:1494–8.

[5] 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.

[6] Akuta N, Suzuki F, Kobayashi M, et al. The influence of hepatitis B virus genotype on the development of lamivudine resistance during long-term treatment. J Hepatol 2003;38:315–21.

[7] 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.

[8] 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.

[9] Kimbi GC, Kramvis A, Kew MC. Distinctive sequence characteristics of subgenotype A1 isolates of hepatitis B virus from South Africa. J Gen Virol 2004;85:1211–20.

[10] 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.

[11] Kew MC, Kramvis A, Yu MC, et al. Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu-speaking subsaharan Africans. J Med Virol 2005;75:513–21.

[12] 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.

[13] Chan HL, Tsui SK, Tse CH, et al. Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. J Infect Dis 2005;191:2022–32.

[14] Usuda S, Okamoto H, Tanaka T, et al. Differentiation of hepatitis B virus genotypes D and E by ELISA using monoclonal antibodies to epitopes on the preS2-region product. J Virol Methods 2000;87:81–9.

[15] Usuda S, Okamoto H, Iwanari H, et al. Serological detection of hepatitis B virus genotypes by ELISA with monoclonal antibodies to type-specific epitopes in the preS2-region product. J Virol Methods 1999;80:97–112.
[16] Mizokami M, Nakano T, Orito E, et al. Hepatitis B virus genotype assignment using restriction fragment length polymorphism patterns. FEBS Lett 1999;450:66–71.
[17] Sugauchi F, Mizokami M, Orito E, et al. A novel variant genotype C of hepatitis B virus identified in isolates from Australian Aborigines: complete genome sequence and phylogenetic relatedness. J Gen Virol 2001;82:883–92.
[18] Kao JH, Chen PJ, Lai MY, et al. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. Gastroenterology 2000;118:554–9.
[19] 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.
[20] Kramvis A, Kew M, Francois G, Hepatitis. B virus genotypes. Vaccine 2005;23:2409–23.
[21] Yuan TT, Faruqi A, Shih JW, et al. The mechanism of natural occurrence of two closely linked HBV precore predominant mutations. Virology 1995;211:144–56.
[22] Li JS, Tong SP, Wen YM, et al. Hepatitis B virus genotype A rarely circulates as an HBe-minus mutant: possible contribution of a single nucleotide in the precore region. J Virol 1993;67: 5402–10.
[23] Chen HS, Kew MC, Hornbuckle WE, et al. The precore gene of the woodchuck hepatitis virus genome is not essential for viral replication in the natural host. J Virol 1992;66: 5682–4.
[24] Sakurai M, Sugauchi F, Tsai N, et al. Genotype and phylogenetic characterization of hepatitis B virus among multi-ethnic cohort in Hawaii. World J Gastroenterol 2004;10:2218–22.
[25] Neurath AR, Kent SB, Strick N, et al. Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. Cell 1986;46:429–36.
[26] Melegari M, Bruno S, Wands JR. Properties of hepatitis B virus pre-S1 deletion mutants. Virology 1994;199:292–300.

Original Paper

Intervirology

Intervirology 839
DOI: 10.1159/0000XXXXX

Received: August 8, 2005
Accepted after revision: August 18, 2005

Variations in the Viral NS5B Region in Japanese Patients with Chronic Hepatitis C Virus Genotype 1b Infection

No Specific Amino Acid Substitution Was Identified as Determinants of Treatment Response to Interferon/Ribavirin Combination Therapy

Kanji Sugihara[a, b] Etsuro Orito[b] Yasuhito Tanaka[a] Takanobu Kato[a] Johnson Y.N. Lau[a] Tomoyoshi Ohno[b] Katsuo Hayashi[c] Masataka Ogino[c] Noboru Hirashima[d] Kenji Sakakibara[d] Yoshiki Mizuno[e] Hideaki Kato[f] Seiji Suzuki[g] Ryuzo Ueda[b] Masashi Mizokami[a]

Departments of [a]Clinical Molecular Informative Medicine and [b]Internal Medicine and Molecular Science, Nagoya City University Graduate School of Medical Sciences, [c]Department of Gastroenterology, Nagoya Daini Red Cross Hospital, [d]Department of Gastroenterology, Chukyo Hospital, and [e]Department of Gastroenterology, Nagoya Higashi General Hospital, Nagoya, [f]Department of Gastroenterology, Toyokawa Municipal Hospital, Toyokawa, and [g]Department of Gastroenterology, Narita Memorial Hospital, Toyohashi, Japan

Key Words
Hepatitis C virus · Nucleotide · Amino acids, substitutions · Interferon · Ribavirin · Genotype

Abstract
Objective: A recent study suggested that the substitution of amino acid 415 of HCV NS5B from phenylalanine to tyrosine in patients with HCV genotype 1a infection is induced by ribavirin and responsible for resistance to ribavirin therapy. The aim of this study was to evaluate whether specific variations in the HCV NS5B sequence in Japanese patients with HCV genotype 1b (HCV/1b) infection are associated with treatment response or selected by treatment with interferon-α/ribavirin combination therapy. ***Methods:*** Eighteen Japanese patients with HCV/1b infection receiving interferon-α/ribavirin combination therapy for 24 weeks were studied. Five patients treated with interferon-α monotherapy for 24 weeks were also studied as controls. The entire HCV NS5B sequence before and after therapy was determined. ***Results:*** All HCV isolates had tyrosine at position 415 of NS5B before and after therapy. Further analysis showed that no specific amino acid substitutions were identified to associate with clinical response and no specific amino acid substitutions were induced/selected by the clinical treatment. ***Conclusion:*** No specific HCV NS5B nucleotide/amino acid sequence variations, including amino acid 415 of NS5B, were identified as being associated with clinical treatment response or selected by the combination therapy in Japanese patients with HCV/1b infection.

The sequences reported in this paper have been deposited in the GenBank/DDBJ/EMBL databases (accession numbers AB189078–AB189119).

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0300–5526/06/0000–0000$23.50/0

Accessible online at:
www.karger.com/int

Masashi Mizokami, MD, PhD
Department of Clinical Molecular Informative Medicine
Nagoya City University Graduate School of Medical Sciences
Nagoya (Japan)
Tel. +81 52 853 8292, Fax +81 52 842 0021, E-Mail mizokami@med.nagoya-cu.ac.jp

Introduction

Hepatitis C virus (HCV) is a single-strand RNA virus of approximately 9,500 nucleotides and is a major etiology of parenteral non-A non-B hepatitis [1, 2]. Natural history studies have shown that a significant proportion of patients with chronic hepatitis C will eventually develop cirrhosis and hepatocellular carcinoma [3]. In Western countries, HCV-related end-stage liver disease is a major indication for liver transplantation [4]. In 1998, the United States Food and Drug Administration approved the interferon-α/ribavirin combination therapy for patients with chronic hepatitis C. A different treatment regimen of interferon-α/ribavirin combination therapy was also approved by the Japanese Health Authority in 2002.

Ribavirin, a synthetic guanosine analog, is an antiviral drug approved for the treatment of respiratory syncytial virus [5] and, in combination with interferon-α, for the treatment of clinically compensated chronic hepatitis C. Ribavirin monotherapy has been shown to reduce serum alanine transaminase (ALT) levels, but no significant reduction in serum HCV RNA levels was observed in most treated patients. Most patients who responded to ribavirin monotherapy relapsed biochemically after cessation of therapy [6]. Subsequently, it was shown that the combination of interferon-α and ribavirin has a much better efficacy than interferon-α monotherapy which was the gold standard of therapy for patients with chronic hepatitis C in the late 1990s [7–9]. However, in patients with HCV genotype-1 infection and high viral load, the response rate to interferon-α/ribavirin combination therapy is still at the 20–30% level, in contrast to a much higher response rate of 80% in patients infected with HCV genotypes 2 and 3 [10].

The non-structural (NS) genomic region of the HCV genome encodes the viral RNA-dependent RNA polymerase, an essential viral replicating enzyme [11]. A recent study suggested that ribavirin treatment in HCV genotype 1a infection might exert selective pressure in favor of a HCV variant with tyrosine (Y) instead of phenylalanine (F) at the 415 position of NS5B. The investigators suggested that this amino acid substitution may lead to viral resistance to ribavirin treatment, as corroborated by their in vitro studies [12].

In Japan, most patients had HCV genotype 1b infection with a high viral titer. The aim of the present study was to evaluate whether variations in the HCV NS5B region play a role in determining the clinical response to interferon-α/ribavirin therapy in Japanese patients with HCV genotype 1b infection.

Patients and Methods

In a clinical study comparing interferon-α/ribavirin combination therapy with interferon-α monotherapy, 70 Japanese patients with HCV genotype 1b infection were recruited in our centers, including 64 patients receiving the combination therapy and 6 patients receiving interferon-α monotherapy. All patients gave written informed consent. All patients were seropositive for anti-HCV and HCV RNA, and genotyping showed HCV genotype 1b infection in all patients. They were all seronegative for HBsAg and anti-HIV, and other causes of liver disease were excluded using standard clinical and laboratory criteria. All patients received treatment in Nagoya City University Hospital and its affiliated hospitals from January 2001 to March 2003. The study was conducted in accordance with the ethical guidelines of the Declaration of Helsinki. Treatment regimen consisted of intramuscular injection of interferon-α2b (Schering KK, Osaka, Japan) at a dose of 6 million units (MU) daily for 2 weeks, followed thrice weekly injection for 22 weeks, combined with oral ribavirin 600–800 mg/day (Schering KK) based on body weight for patients with combination therapy. Patients were followed for an additional 24 weeks after cessation of therapy to determine the long-term clinical response. The 24-week treatment protocol was approved by the Japanese Health Authority back in 1999. More recently, the Japanese Health Authority has approved studies based on 48 weeks of treatment.

Clinical treatment response was classified into (1) sustained virologic response (SVR), defined as undetectable serum HCV RNA and normal ALT at week 24 after cessation of the treatment; (2) relapsers (Rel), defined as converting to negative for serum HCV RNA during therapy but relapsed to positive serum HCV RNA within 24 weeks after cessation of treatment, and (3) non-responders (NR), defined as patients who retained serum HCV RNA at the end of treatment. Of the 64 patients who received interferon-α/ribavirin combination therapy, only 45 completed the treatment without a dose reduction/discontinuation, the others had a dose reduction or treatment discontinuation due to side effects. Of the 45 patients who completed their therapy, a SVR was observed in 7 patients (15.6%), 25 patients were Rel (55.6%), and 13 patients were NR (28.9%). There were no significant differences in gender distribution, age, pretreatment ALT level, serum HCV RNA level, blood hemoglobin level, and platelet count between patients in the SVR, Rel and NR groups.

Eighteen patients with different clinical response profiles (SVR 4, Rel 8, NR 6) were randomly selected for this study. Five of the 6 patients who received interferon-α were also included to serve as controls (SVR 1, Rel 1, NR 3). As the Japanese Health Authority did not authorize the ribavirin monotherapy arm in any clinical protocol due to a poor benefit to risk ratio, there was no access to ribavirin monotherapy for Japanese patients as controls. Serum samples before and after therapy were used for this study. The serum samples were collected and serum separated from clots within 3 h and stored at –80°C.

All patients gave written informed consent and this study was approved by the local institutional review board.

Methods

Nucleic acids were extracted using a SepaGean RV-R Nucleic acid extraction kit (Sanko Junyaku Co., Ltd., Tokyo, Japan) in accordance with the manufacturer's protocol. The extracted RNA was reverse-transcribed to cDNA using SuperScript II RNase H⁻

Table 1. Background of the patients studied

Patient No.	Clinical treatment response	Gender	HCV genotype	Age years	Pretreatment		
					RNA level KIU/ml	ALT IU/l	Hb mg/dl
M1	SVR	M	1b	43	>850	361	14.0
M2	Rel	M	1b	53	560	54	14.6
M3	NR	M	1b	42	700	212	14.5
M4	NR	F	1b	44	620	209	13.8
M5	NR	M	1b	51	>850	108	13.6
C1	SVR	M	1b	54	830	112	15.1
C2	SVR	F	1b	60	>850	51	14.2
C3	SVR	M	1b	35	>850	134	16.6
C4	SVR	F	1b	54	>850	26	12.6
C5	Rel	M	1b	32	410	51	14.2
C6	Rel	M	1b	67	190	154	13.9
C7	Rel	M	1b	68	460	164	13.0
C8	Rel	M	1b	68	>850	98	12.7
C9	Rel	M	1b	47	560	50	15.6
C10	Rel	F	1b	64	>850	16	13.7
C11	Rel	M	1b	44	>850	56	15.5
C12	Rel	F	1b	70	>850	142	12.2
C13	NR	M	1b	61	>850	61	13.5
C14	NR	M	1b	67	260	139	13.1
C15	NR	M	1b	52	>850	113	11.7
C16	NR	M	1b	70	400	101	16.7
C17	NR	M	1b	29	>850	55	14.1
C18	NR	F	1b	60	>850	90	14.3
Mean ± SD				54 ± 13	697 ± 217	111 ± 77	14.1 ± 1.3

M1–5 = Patients treated with interferon-α monotherapy; C1–18 = patients treated with interferon-α/ribavirin combination therapy; ALT = serum alanine transaminase; Hb = hemoglobin.

Reverse Transcriptase (Invitrogen Corp., Carlsbad, Calif., USA) and random hexamer primer (Takara Shuzo Co. Ltd, Tokyo, Japan) as described previously [13].

The complete HCV NS5B sequence was determined in all samples. Briefly, cDNA was amplified by a long polymerase chain reaction (PCR) with LA Taq (Takara Shuzo Co. Ltd, Tokyo, Japan). The long PCR fragments (around 2.1 kbp) were generated with primers 7355F-NK (CCTGACAGAGTCCACCGTGTCTTCTG CCTT) and 9440R-NK (GTTGGGGAGCAGGTAGATGCCT ACCCCTAC). The first-round PCR product was further amplified using hemi-nested primers including 7378F-NK (AGCTCGC-TACTAAGACCTTTGGCAGCTCCG) and 9440R-NK. The amplicons were then cloned into pCR2.1-TOPO vector (Invitrogen Corp), and nucleotide sequences were determined using Prism Big Dye (Applied Biosystems, Foster City, Calif., USA) with an ABI 3100 DNA automated sequencer.

The sequences generated were used to confirm HCV genotypes and to identify specific nucleotide mutations and amino acid substitutions that may be associated with interferon-α/ribavirin treatment. Additionally, to determine whether ribavirin serves as a RNA mutagen, two models were used to estimate evolutionary distances: Tamura-Nei model for all codon substitutions, and Pamilo-Bianchi-Li model for synonymous and non-synonymous substitutions.

Amino acid sequences of the HCV NS5B region in various HCV genotypes were obtained from GenBank/DDBJ database. They were aligned to show amino acid 415 of the HCV NS5B region.

Statistical Analysis

The χ^2 test, Fisher's exact test and Student's t test were used where appropriate.

Results

Nucleotide Substitutions in the HCV NS5B Region before and after Therapy

The mean pair-wise genetic distances of all codon substitutions among 14 patients (4/18 had a SVR and were negative during follow-up) receiving combination therapy were 0.00668 ± 0.00702, which was not different from

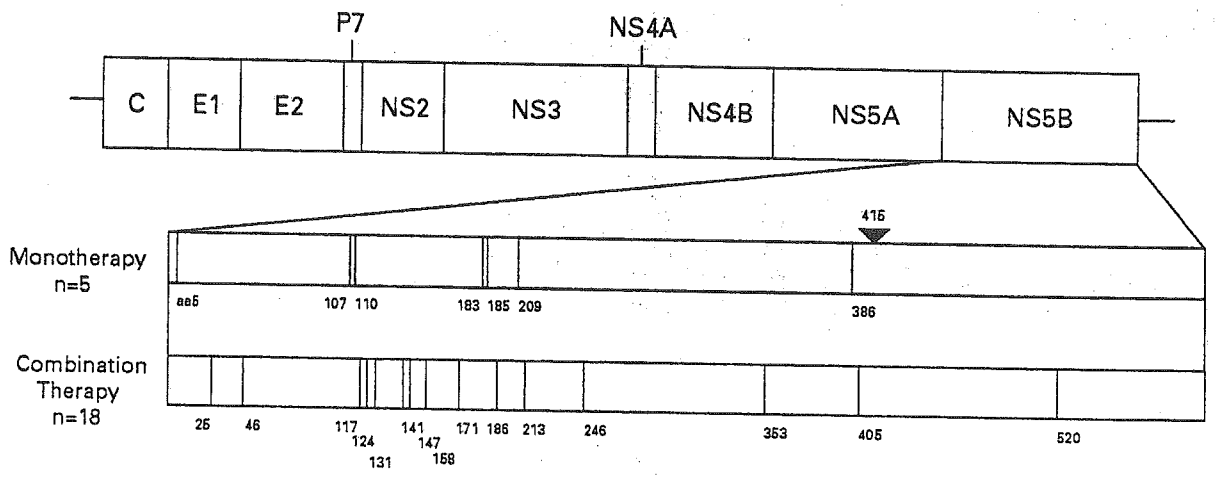

Fig. 1. The site map of amino acid mutations in the NS5B region in patients receiving interferon-α monotherapy versus interferon-α/ribavirin combination therapy. The vertical bars indicate the sites of amino acid substitutions after therapy.

the 4 patients (1/5 had a SVR) receiving interferon-α monotherapy 0.00627 ± 0.00444 (p = NS). The nucleotide substitution rate was also not different between the 2 groups after taking account of the bias towards synonymous substitutions (0.01778 ± 0.0237 vs. 0.01603 ± 0.01208, p = NS).

Amino Acid Substitutions in the NS5B Region

In the 5 patients receiving interferon-α monotherapy (1 was SVR and negative for HCV RNA during follow-up), 7 amino acid substitutions in the NS5B region were detected. In the 18 patients treated with interferon-α/ribavirin combination therapy (4 were SVR and negative for HCV RNA during follow-up), 17 amino acid substitutions were detected (fig. 1; table 2–4). There were no amino acid substitutions specific to the clinical treatment response identified. Note that with reference to the amino acid residue at the NS5B 415th site, all 18 in 23 (since 5/23 were SVRs) patients with HCV genotype 1b infection had Y before and after treatment.

There was also no difference in the number of sites of amino acid substitutions in the HCV NS5B region between 17 in the patients treated with combination therapy and 7 in the patients receiving interferon-α monotherapy (0.2055% (17/591 × 14) vs. 0.2961% (7/591 × 4); the NS5B region consists of 591 amino acids).

Amino Acid Alignment in the NS5B Region of Different HCV Genotypes

A total of 150 complete amino acid sequences derived from the nucleotide sequences of the HCV NS5B region from different HCV genotypes (from genotype 1a to 6) were obtained from the GenBank/DDBJ database. All 17 HCV-1a sequences had F at position 415 of the NS5 region. In contrast, 104/106 HCV genotype 1b sequences had Y at position 415 of NS5B. For other HCV genotypes, only HCV genotype 1c and 1 of 6 genotype 3 sequences had F at position 415 of the NS5B region (fig. 2; table 5).

Discussion

This study showed three important points. First, most HCV genotypes 1b had Y, instead of F in position 415 of the HCV NS5B region. Second, interferon-α/ribavirin treatment did not induce any specific amino acid substitutions in the HCV NS5B region in Japanese patients with HCV genotype 1b infection. Third, interferon-α/ribavirin treatment did not induce more nucleotide mutations or amino acid substitutions compared to interferon-α monotherapy in Japanese patients with HCV genotype 1b infection.

Table 2. Amino acid substitutions in HCV NS5B in the patients studied

NS5B	aa	5th		25th		46th		107th		110th		117th		124th		131th	
		before	after	before	after	before	after	before	after	before	after	before	after	before	after	before	after
M1	SVR	T	ND	A	ND	S	ND	D	ND	N	ND	N	ND	K	ND	E	ND
M2	Rel	T	T	A	A	S	S	D	D	N	N	N	N	K	K	E	E
M3	NR	T	T	A	A	S	S	D	D	N	N	N	N	K	K	E	E
M4	NR	S	T	P	P	G	G	D	N	S	N	N	N	K	K	E	E
M5	NR	T	T	A	A	S	S	D	D	N	N	N	N	K	K	E	E
C1	SVR	T	ND	P	ND	G	ND	D	ND	N	ND	N	ND	K	ND	D	ND
C2	SVR	T	ND	A	ND	S	ND	D	ND	N	ND	N	ND	K	ND	D	ND
C3	SVR	T	ND	A	ND	G	ND	D	ND	S	ND	N	ND	K	ND	E	ND
C4	SVR	T	ND	A	ND	G	ND	D	ND	S	ND	N	ND	K	ND	E	ND
C5	Rel	T	T	A	A	G	G	D	D	N	N	N	N	K	K	D	D
C6	Rel	T	T	P	P	S	S	D	D	S	S	N	N	K	K	E	E
C7	Rel	T	T	A	A	S	S	D	D	N	N	N	N	K	K	E	E
C8	Rel	T	T	P	P	S	S	D	D	N	N	N	N	E	E	E	E
C9	Rel	T	T	P	P	S	S	D	D	N	N	N	N	K	K	E	E
C10	Rel	T	T	A	S	C	S	D	D	N	N	N	N	K	K	E	E
C11	Rel	T	T	A	A	S	S	D	D	N	N	N	N	K	K	E	E
C12	Rel	T	T	A	A	G	G	D	D	S	S	N	N	K	K	E	E
C13	NR	T	T	P	A	S	S	D	D	N	N	D	N	K	E	E	E
C14	NR	T	T	A	P	S	S	D	D	N	N	N	N	E	E	E	E
C15	NR	T	T	A	A	S	S	D	D	N	N	N	N	K	K	V	V
C16	NR	T	T	A	A	S	S	D	D	S	S	N	N	K	K	E	E
C17	NR	T	T	A	A	S	S	D	D	N	N	N	N	K	K	E	E
C18	NR	T	T	A	A	S	S	D	D	N	N	N	N	K	K	D	E

ND = Not detected; M1–5 = patients treated with interferon-α monotherapy; C1–18 = patients treated with interferon-α/ribavirin combination therapy.

Table 3. Amino acid substitutions in HCV NS5B in the patients studied

NS5B	aa	141th		147th		158th		171th		183th		185th		186th		209th	
		before	after	before	after	before	after	before	after	before	after	before	after	before	after	before	after
M1	SVR	K	ND	V	ND	R	ND	E	ND	P	ND	A	ND	V	ND	K	ND
M2	Rel	K	K	V	V	R	R	E	E	P	L	A	A	V	V	K	K
M3	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
M4	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
M5	NR	K	K	V	V	R	R	E	E	P	P	A	V	V	V	R	K
C1	SVR	K	ND	V	ND	R	ND	E	ND	P	ND	A	ND	V	ND	K	ND
C2	SVR	K	ND	V	ND	R	ND	E	ND	P	ND	A	ND	V	ND	K	ND
C3	SVR	K	ND	I	ND	R	ND	E	ND	P	ND	A	ND	V	ND	K	ND
C4	SVR	K	ND	I	ND	R	ND	E	ND	P	ND	A	ND	V	ND	K	ND
C5	Rel	K	V	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C6	Rel	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C7	Rel	E	K	V	I	R	R	E	E	P	P	A	A	V	V	K	K
C8	Rel	K	K	I	I	R	R	E	E	P	P	A	A	V	V	K	K
C9	Rel	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C10	Rel	K	K	I	I	R	R	E	E	P	P	A	A	V	V	K	K
C11	Rel	K	K	V	V	R	R	E	E	P	P	A	A	G	G	K	K
C12	Rel	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C13	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C14	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C15	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C16	NR	K	K	V	V	R	G	E	K	P	P	A	A	V	G	K	K
C17	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K
C18	NR	K	K	V	V	R	R	E	E	P	P	A	A	V	V	K	K

ND = Not detected; M1–5 = patients treated with interferon-α monotherapy; C1–18 = patients treated with interferon-α/ribavirin combination therapy.

Table 4. Amino acid substitutions in HCV NS5B in the patients studied

NS5B	aa	213th		246th		353th		386th		405th		415th		520th	
		before	after	before	after	before	after	before	after	before	after	before	after	before	after
M1	SVR	C	ND	A	ND	P	ND	R	ND	V	ND	Y	ND	T	ND
M2	Rel	C	C	A	A	P	P	R	C	V	V	Y	Y	T	T
M3	NR	C	C	A	A	P	P	R	R	V	V	Y	Y	T	T
M4	NR	C	C	A	A	L	L	R	R	V	V	Y	Y	T	T
M5	NR	C	C	A	A	P	P	R	R	V	V	Y	Y	T	T
C1	SVR	C	ND	A	ND	L	ND	R	ND	V	ND	Y	ND	T	ND
C2	SVR	T	ND	A	ND	P	ND	R	ND	V	ND	Y	ND	T	ND
C3	SVR	N	ND	A	ND	P	ND	R	ND	V	ND	Y	ND	T	ND
C4	SVR	C	ND	A	ND	P	ND	R	ND	V	ND	Y	ND	T	ND
C5	Rel	N	N	A	V	P	P	R	R	V	V	Y	Y	T	T
C6	Rel	C	C	A	A	L	L	R	R	V	V	Y	Y	T	I
C7	Rel	N	N	A	A	P	L	R	R	V	V	Y	Y	T	T
C8	Rel	C	C	A	A	L	L	R	R	V	V	Y	Y	T	T
C9	Rel	N	N	A	A	P	P	R	R	V	V	Y	Y	T	T
C10	Rel	N	N	A	A	L	L	R	R	V	V	Y	Y	T	T
C11	Rel	S	S	A	A	P	P	R	R	V	V	Y	Y	T	T
C12	Rel	C	C	A	A	P	P	R	R	V	V	Y	Y	T	T
C13	NR	S	N	A	A	P	P	R	R	V	V	Y	Y	T	T
C14	NR	C	C	A	A	P	P	R	R	V	V	Y	Y	T	T
C15	NR	T	T	A	A	P	P	R	R	V	V	Y	Y	T	T
C16	NR	R	C	A	A	P	P	R	R	I	V	Y	Y	T	T
C17	NR	C	C	A	A	P	P	R	R	V	V	Y	Y	T	T
C18	NR	C	C	A	A	P	P	R	R	V	V	Y	Y	T	T

ND = Not detected; M1–5 = patients treated with interferon-α monotherapy; C1–18 = patients treated with interferon-α/ribavirin combination therapy.

Table 5. Amino acid residues at position 415 of NS5B in various HCV genotypes

HCV genotype	No.	Position 415 of NS5B			
		F	%	Y	%
1a	17	17	100	0	0
1b	106	2	2	104	98
1c	2	2	100	0	0
2a	18	0	0	18	100
2b	3	0	0	3	100
3	6	1	17	5	83
4	1	0	0	1	100
5	2	0	0	2	100
6	5	0	0	5	100

Understanding the molecular mechanism(s) of action of interferon-α/ribavirin therapy and determinants/resistance to therapy may lead to the design of better treatment strategies. The currently proposed mechanisms of action of ribavirin in combination therapy with interferon-α for HCV include: (1) inducing a Th2 to Th1 bias in favor of a host antiviral response; (2) blocking the host enzyme inosine monophosphate dehydrogenase to reduce the availability of the guanosine pool; (3) inhibition of viral RNA-dependent RNA polymerase, and (4) serving as a RNA mutagen to introduce mutations into the HCV genome [14]. Recently, it was observed that a HCV variant with amino acid substitution from F to Y at the NS5B 415th position was consistently detected in patients with HCV genotype 1a infection during ribavirin monotherapy. It was suggested that this amino acid substitution was selected by ribavirin therapy and was responsible for viral resistance during therapy [12].

This study showed that amino acid substitution at the NS5B 415 position is irrelevant in Japanese patients with HCV genotype 1b infection. First, all the Japanese patients with HCV genotype 1b that we studied had Y at the 415 position of the NS5B region. Second, this is confirmed by the HCV gene and deduced amino acid sequences available in the Genbank/DDBJ database. Such a pattern with F at the 415 position of HCV NS5B region was commonly observed in HCV genotypes 1a and 1c, and only occasionally in genotypes 1b and 3. To argue that F at position 415 is an important phenotype as a re-

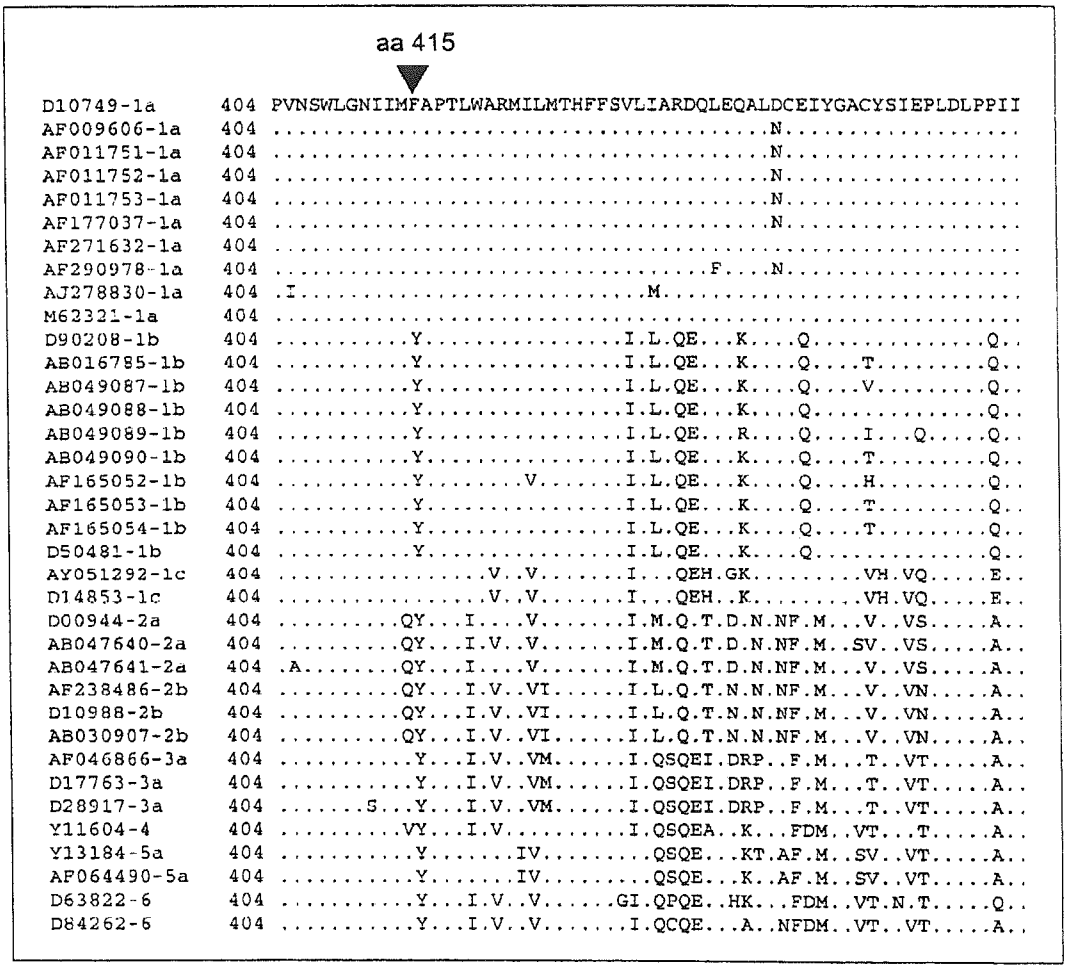

Fig. 2. The alignment of the deduced amino acid sequences of various HCV genotypes in the NS5B region based on 150 sequences available in the GenBank/DDBJ database.

sponse to therapy determinant is difficult since most isolates of HCV genotypes 2 and 3, which are known to respond very well to combination therapy, have Y at position 415 of NS5B. One has to postulate that there are other significant viral or host factors related to HCV genotypes 2 and 3 that make them more susceptible to response to combination therapy to offset the Y factors at position 415 of HCV NS5B. Also, the clinical observation that HCV genotype 1a (with F in position 415) and genotype 1b (with Y in position 415) had a similar clinical treatment response rate to interferon-α/ribavirin therapy also suggests that the amino acid in position 415 is not a key determinant to clinical response to interferon-α/ribavirin combination therapy.

The entire NS5B region was studied before and after therapy and no specific amino acid substitutions were found to be associated with relapse or no response. Therefore, this study ruled out the possibility of a viral amino acid substitution in the NS5B region as a major viral determinant for response to therapy or a viral resistance factor.

Finally, ribavirin has been suggested to be a viral mutagen. In this study, the number of nucleotide mutations and amino acid substitutions in the NS5B region did not increase with interferon-α/ribavirin combination therapy compared with interferon-α monotherapy. Certainly, the best approach is to study ribavirin monotherapy in Japanese patients with HCV genotype 1b infection. However,

the risk to benefit ratio precludes this type of therapy for testing in Japan. Nevertheless, the present study showed that the mutagen effect of ribavirin was not observed in Japanese patients with genotype 1b infection after a 6-month course of interferon-α/ribavirin combination therapy.

Acknowledgements

The authors thank Dr. Hiroshi Tokuda, Dr. Kaoru Suzuki, Dr. Makoto Nakamura, Dr. Makoto Narita, and Dr. Yuko Matsumoto for their support of this study, and Mrs. Kyoko Akita and Yoshiko Kobayashi for their editorial assistance. This study was supported in part by grants from the Ministry of Health, Labor and Welfare of Japan, and grants-in-aid for Young Scientists (A) from the Ministry of Education, Culture, Science, and Sports of Japan (16689016).

References

1 Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M: Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 1989; 244:359–362.

2 Major ME, Feinstone SM: The molecular virology of hepatitis C. Hepatology 1997;25: 1527–1538.

3 Kiyosawa K, Sodeyama T, Tanaka E, Gibo Y, Yoshizawa K, Nakano Y, Furuta S, Akahane Y, Nishioka K, Purcell RH, Alter HJ: Interrelationship of blood transfusion, non-A, non-B hepatitis and hepatocellular carcinoma: analysis by detection of antibody to hepatitis C virus. Hepatology 1990;12:671–675.

4 National Institutes of Health Consensus Development Conference Statement: Management of hepatitis C: 2002–June 10–12, 2002. Hepatology 2002;36(suppl 1):S3–20.

5 Wyde PR: Respiratory syncytial virus (RSV) disease and prospects for its control. Antiviral Res 1998;39:63–79.

6 Kakumu S, Yoshioka K, Wakita T, Ishikawa T, Takayanagi M, Higashi Y: A pilot study of ribavirin and interferon beta for the treatment of chronic hepatitis C. Gastroenterology 1993; 105:507–512.

7 McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, Goodman ZD, Ling MH, Cort S, Albrecht JK: Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. N Engl J Med 1998;339:1485–1492.

8 Brillanti S, Garson J, Foli M, Whitby K, Deaville R, Masci C, Miglioli M, Barbara L: A pilot study of combination therapy with ribavirin plus interferon alfa for interferon alfa-resistant chronic hepatitis C. Gastroenterology 1994;107:812–817.

9 Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, Bain V, Heathcote J, Zeuzem S, Trepo C, Albrecht J: Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. International Hepatitis Interventional Therapy Group (IHIT). Lancet 1998;352:1426–1432.

10 Davis GL, Esteban-Mur R, Rustgi V, Hoefs J, Gordon SC, Trepo C, Shiffman ML, Zeuzem S, Craxi A, Ling MH, Albrecht J: Interferon alfa-2b alone or in combination with ribavirin for the treatment of relapse of chronic hepatitis C. International Hepatitis Interventional Therapy Group. N Engl J Med 1998;339: 1493–1499.

11 Behrens SE, Tomei L, De Francesco R: Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. EMBO J 1996;15:12–22.

12 Young KC, Lindsay KL, Lee KJ, Liu WC, He JW, Milstein SL, Lai MM: Identification of a ribavirin-resistant NS5B mutation of hepatitis C virus during ribavirin monotherapy. Hepatology 2003;38:869–878.

13 Ohno T, Mizokami M, Wu RR, Saleh MG, Ohba K, Orito E, Mukaide M, Williams R, Lau JY: New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. J Clin Microbiol 1997;35:201–207.

14 Lau JY, Tam RC, Liang TJ, Hong Z: Mechanism of action of ribavirin in the combination treatment of chronic HCV infection. Hepatology 2002;35:1002–1009.

tapraid2/zg0-gast/zg0-gast/zg000306/zg02624d06z | xppws | S=1 | 1/27/06 | Art: 2624 | DTD5.0

GASTROENTEROLOGY 2006;xx:xxx

Molecular Tracing of the Global Hepatitis C Virus Epidemic Predicts Regional Patterns of Hepatocellular Carcinoma Mortality

YASUHITO TANAKA,* FUAT KURBANOV, SHUHEI MANO,† ETSURO ORITO,§ VICTOR VARGAS,¶ JUAN I. ESTEBAN,¶ MAN-FUNG YUEN,‖ CHING-LUNG LAI,‖ ANNA KRAMVIS,# MICHAEL C. KEW,# HEIDI E. SMUTS,** SERGEY V. NETESOV,†† HARVEY J. ALTER,§§ and MASASHI MIZOKAMI*

*Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; †Nagoya City University Graduate School of Natural Sciences, Nagoya, Japan; §Department of Internal Medicine and Molecular Science, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; ¶Department of Internal Medicine, Hospital General Universitari Vall d'Hebron, Barcelona, Spain; ‖Department of Medicine, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China; #MRC/CANSA/University Molecular Hepatology Research Unit, Department of Medicine, University of the Witwatersrand, Johannesburg, South Africa; **Division of Medical Virology/National Health Laboratory Services, University of Cape Town, Cape Town, South Africa; ††State Research Center of Virology and Biotechnology VECTOR Koltsovo, Novosibirsk Region, Russian Federation; §§Department of Transfusion Medicine, Warren Grant Magnuson Clinical Center, The National Institutes of Health, Bethesda, Maryland USA

Background & Aims: Molecular evolutionary analysis based on coalescent theory can provide important insights into epidemiologic processes worldwide. This approach was combined with analyses of the hepatitis C virus (HCV) epidemiologic-historical background and HCV-related hepatocellular carcinoma (HCC) in different countries. *Methods:* The HCV gene sequences of 131 genotype 1b (HCV-1b) strains from Japan, 38 HCV-1a strains from the United States, 33 HCV-1b strains from Spain, 27 HCV-3a strains from the former Soviet Union (FSU), 47 HCV-4a strains from Egypt, 25 HCV-5a strains from South Africa, and 24 HCV-6a strains from Hong Kong isolated in this study and previous studies were analyzed. *Results:* The coalescent analysis indicated that a transition from constant size to rapid exponential growth (spread time) occurred in Japan in the 1920s (HCV-1b), but not until the 1940s for the same genotype in Spain and other European countries. The spread time of HCV-1a in the United States was estimated to be in the 1960s, HCV-3a in the FSU, HCV-5a in South Africa, and HCV-6a in Hong Kong in the 1960s, mid-1950s, and late 1970s, respectively. Three different linear progression curves were determined by analysis of the relationship between HCV seroprevalence and HCC mortality in different geographic regions; a steep ascent indicated the greatest progression to HCC in Japan, a near horizontal line indicated the least progression in the United States and the FSU, and an intermediate slope was observed in Europe. *Conclusions:* These findings strongly suggest that the initial spread time of HCV is associated with the progression dynamics of HCC in each area, irrespective of genotype.

Chronic hepatitis C virus (HCV) infection is an endemic disease affecting millions of individuals worldwide.[1,2] HCV infection usually is clinically mild, but the stages of more than 20% of patients can progress during the clinical course, occasionally culminating in hepatocellular carcinoma (HCC) over the course of 2–3 decades, the latter especially in Japan, Spain, and Italy.[3–6] Because the time lag between HCV infection and cancer development is several decades,[4] it is important to estimate the demographic history of HCV infection to predict the future burden of disease.

HCV is classified into 6 major genotypes.[7–9] Within the genotypes there are many subtypes, with varying geographic distributions and modes of transmission.[7] Subtypes 1a (HCV-1a), 1b, 2a, 2b, and 3a are distributed globally and account for the majority of HCV infections worldwide.[10,11] The rapid spread and global dissemination of these subtypes arises from their efficient parenteral transmission via transfusion of contaminated blood products, medical procedures, and illegal injection drug use. Other endemic and epidemic HCV strains are found in restricted geographic areas, including HCV-4a in Egypt, 5a in South Africa, and 6a in southeast Asia (Hong Kong).[7] Because HCV was not identified until 1989, it is difficult to estimate epidemic dynamics associated with the subtypes prevalent before this time.

Abbreviations used in this paper: FSU, former Soviet Union; IDU, injection drug use; *Sj*, *Schistosoma japonicum*; SRDT, single rate-dated tips.

0016-5085/06/$32.00
doi:10.1053/j.gastro.2006.01.032

tapraid2/zg0-gast/zg0-gast/zg000306/zg02624d06z | xppws | S=1 | 1/27/06 | Art: 2624 | DTD5.0

Table 1. Characteristics of the Population in Each Country

	United States	Japan *Sj* group	Japan non-*Sj*	Spain	FSU	Egypt	South Africa	Hong Kong
Number	38[a]	64[a]	67[a]	33	27[a]	47[a]	25	24
Mean age	48.2 ± 11.7	69.9 ± 7.7	67.2 ± 8.8	52.9 ± 11.0	24.7 ± 3.6	38.8 ± 9.0	56.8 ± 10.5	49.1 ± 15.4
Sex (M/F)	24/14	34/30	33/34	17/16	22/5	33/14	17/8	12/12
HCV subtypes	1a	1b	1b	1b	3a	4a	5a	6a
Divergence time[b]	1920	1812	1918	1892	1958	1902	1937	1963
Spread time	1965 (1958–1970)	1923 (1890–1937)	1940 (1933–1948)	1942 (1934–1955)	1963 (1958–1974)	1930 (1917–1940)	1955 (1948–1962)	1977 (1968–1982)
Growth rate, y^{-1}	.15298 (.106181–.213943)	.06667 (.044460–.094732)	.12008 (.098260–.149888)	.09715 (.074642–.140650)	.15625 (.121928–.215722)	.09004 (.074114–.115969)	.19762 (.107049–.356198)	.17120 (.090839–.280186)
Risk factors	IDU	PAT	IDU, transfusion, medical	IDU, transfusion, medical	IDU	PAT	Transfusion, medical	IDU

NOTE. 95% confidence intervals shown in parentheses.

PAT, parenteral antischistosomal therapy; transfusion, blood transfusion; medical, medical procedures.

[a]Eight sequences in the United States, 131 in Japan, 20 in the FSU, and 47 in Egypt were obtained from our previous data.[14–17]

[b]Divergence time indicates the most recent common ancestor (MRCA) point of each subtype.

However, using methods based on coalescent theory,[12] the epidemic history of HCV population can be reconstructed from observed genetic diversity of the viral strains.[13] Recently, the molecular clock theory has been applied successfully to estimate the molecular evolutionary rate in long-term serial serum samples obtained from HCV-infected patients in the United States and Japan; a 30-year lag in HCV spread time was shown between these countries. Insofar as a long duration of HCV infection is a critical determinant for the development of HCC, the molecular clock predicted that the incidence of HCC will increase in the United States over the next 2–3 decades[14] and approach the high rates currently observed in Japan.

In a previous study,[15] the spread of HCV-1a, 1b, 4, and 6 infections worldwide was analyzed by the use of HCV sequences obtained from DNA databases; however, corollary clinical and demographic data were limited. In the present study, new sequences from wider geographic regions and with more extensive clinical information are presented. Specifically, HCV-1a strains in the United States, HCV-1b in Japan,[14,15] HCV-1b in Spain, HCV-3a in the former Soviet Union (FSU),[16,17] HCV-4a in Egypt,[18] HCV-5a in South Africa, and HCV-6a in Hong Kong were analyzed by a coalescent-based approach using principles of both population genetics and mathematic epidemiology.[13] Furthermore, the relationship between the estimated spread time and HCC mortality in each country is discussed.

Materials and Methods

HCV Serum Samples From the United States, Spain, FSU, South Africa, and Hong Kong

To elucidate the epidemic history of HCV population in each country, 30 HCV-1a, 33 HCV-1b, 7 HCV-3a, 25 HCV-5a, and 24 HCV-6a samples were obtained from the following blood banks or hospitals, respectively: National Institutes of Health (United States), Hospital Vall d'Hebron (Spain), National Reference Laboratory of Ministry of Health (Uzbekistan), University of the Witwatersrand and National Health Laboratory Services/University of Cape Town (South Africa), and Queen Mary Hospital (Hong Kong). The samples were collected between 2000 and 2003. Along with our previous data,[14–18] the characteristics of the populations studied are shown in Table 1; the mean age of the Japanese population was the oldest and that of the populations from the FSU and Egypt was significantly younger. The study protocol conformed to the 1975 Declaration of Helsinki, and was approved by the Ethic Committees from each institution. Every patient gave written informed consent to participate in the virologic research of HCV at each blood center or hospital. None of the patients had been treated with interferon therapy for HCV infection.

HCV Gene Sequences

The HCV subtypes studied in each country are shown in Table 1, including 38 HCV-1a sequences from the United States (30 newly determined and 8 previously reported sequences[14]); 131 previously reported HCV-1b sequences including 64 from *Schistosoma japonicum* (*Sj*)-positive sera and 67 from *Sj*-negative (non-Sj) sera in Japan[15]; 33 HCV-1b sequences from Spain; 27 HCV-3a sequences from the FSU (7 newly determined and 20 previously reported sequences[16,17]); 47 previously reported HCV-4a sequences from Egypt[18]; 25 HCV-5a sequences from South Africa; and 24 HCV-6a sequences from Hong Kong. The GenBank/DDBJ accession numbers of the sequences obtained in the present study are AB204592–AB204708. Japanese HCV-1b and Egyptian HCV-4a sequences were obtained from our previous data (AB103424–AB103457, AF271800–AF271812).[15,18] Available related sequences in

tapraid2/zg0-gast/zg0-gast/zg000306/zg02624d06z | xppws | S=1 | 1/27/06 | Art: 2624 | DTD5.0

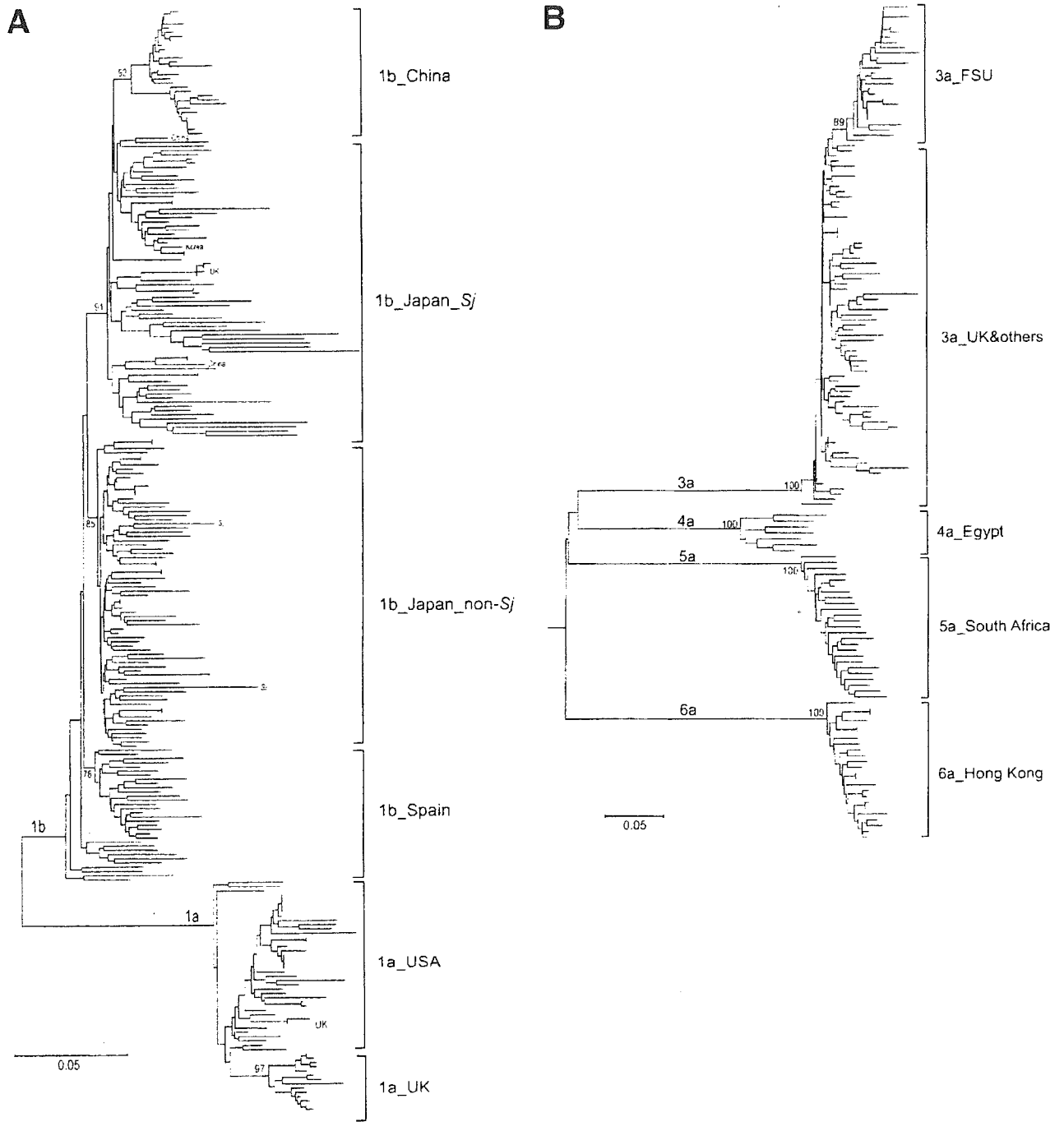

Figure 1. Phylogenetic trees constructed on NS5B sequences of (*A*) HCV-1a and HCV-1b strains and (*B*) HCV-3a, -4a, -5a, and -6a strains. The HCV-1a strains in the United States and HCV-1b strains in Japan (*Sj* and non-*Sj*) and Spain had each significant cluster, and HCV-3a in the FSU, HCV-4a in Egypt, HCV-5a in South Africa, and HCV-6a in Hong Kong. The numbers in the tree indicate bootstrap reliability by the interior branch test. Exceptional strains were labeled according to their country of origin. Significant clusters were subjected to population history analyses.

Figure 1 were recruited from the Hepatitis Virus Database (http://s2as02.genes.nig.ac.jp).

Genotyping and Sequencing

Nucleic acids were extracted from the serum samples using a SepaGene RV-R Nucleic acid extracting kit (Sanko Junyaku Co., Ltd., Tokyo, Japan) in accordance with the manufacturer's protocol. Viral RNA was reverse-transcribed to complementary DNA using SuperScript II RNase H⁻ Reverse Transcriptase (Invitrogen Corp., Carlsbad, CA) and random hexamer primer (Takara Shuzo Co. Ltd, Tokyo, Japan) as described previously.[19] A sequence spanning 339 nucleotides

tapraid2/zg0-gast/zg0-gast/zg000306/zg02624d06z | xppws | S=1 | 1/27/06 | Art: 2624 | DTD5.0

in the NS5B region was amplified by polymerase chain reaction with primers described previously.[14] Polymerase chain reaction products were sequenced directly with Prism Big Dye (Applied Biosystems, Foster City, CA) in an ABI 3100 DNA automated sequencer. To reduce the number of artificial substitutions arising in polymerase chain reaction, Platinum Pfx DNA Polymerase (Invirogen Corp.) with a very high fidelity was used. The sequences determined were used to confirm HCV genotypes and to construct phylogenetic trees. To confirm the reliability of the phylogenetic tree, bootstrap reliability was performed by the interior branch test.[20] The overall mean genetic distances in all nucleotide positions and synonymous and nonsynonymous positions were estimated by MEGA version 2.1.[21]

Analysis of Isolation and Migration of HCV Sequences Among Countries

The phylogeny of the HCV-1a, -1b, and -3a sequences from all of the countries under investigation were estimated by the Neighbor–Joining method[22] using the MEGA software (version 3).[21] The analysis of isolation and migration of the sequences was performed by using a method conducted by Nakano et al.[15] The migration histories of the HCV-1a, -1b, and -3a infections were inferred from the phylogeny by the parsimony method whereby the states are the sampling countries of each sequence and the state changes represent migration events. We calculated the expectations and the statistical significances based on the null distribution generated by 1000 times randomization of the sequences with fixing the topology of the phylogeny, where the null hypothesis is that all of the sequences were sampled from a hypothetic panmictic population, using software we developed ourselves.

Estimating Evolutionary Rates and Dating the Origin of HCV

A reconstructed tree was built on the NS5B sequence of 339 nucleotides by a heuristic maximum-likelihood topology search with stepwise addition and the nearest-neighbor–interchange algorithms. Tree likelihood scores were calculated using the HKY85+G method with the molecular clock enforced by PAUP version 4.0b8. By using the estimated topology, all possible root positions were evaluated under a single rate-dated tips (SRDT) model with the computer software TipDate v1.2 and the root that yielded the highest likelihood was adopted.[23] The program provided a maximum-likelihood estimate of the rate and also the associated date of the most recent common ancestor of the sequences, using a model that assumed a constant rate of nucleotide substitution. The molecular clock was tested by a likelihood ratio test between the SRDT model and a general unconstrained branch length model (different rate model). To confirm the reliability of the phylogenetic tree, bootstrap resampling tests also were performed 1000 times.

Demographic Model

As estimates of the demographic history, a nonparametric function, known also as the *skyline plot*, was obtained by transforming coalescent intervals of an observed genealogy into a piecewise plot that represents an effective number of infections through time.[15,24] A parametric maximum-likelihood was estimated by several models with the computer software Genie v3.5 to build a statistical framework for inferring the demographic history of a population on phylogenies reconstructed on sampled DNA sequences.[25] This model assumes a continuous epidemic process in which the viral transmission parameters remain constant through time. Model fitting was evaluated by likelihood-ratio tests of the parametric maximum-likelihood estimates.[24,25]

Results

Analysis of Isolation and Migration of HCV Sequences Among Countries

Preliminarily, all sequences generated in this study were subjected to phylogenetic analyses together with all previously reported sequences available from DDBJ/GeneBank. The most significant phylogenetic clusters containing a total of 325 representatives of the HCV endemic populations from different regions (Table 1) were determined and subjected to further maximum-likelihood phylogenetic analysis with enforced molecular clock, as previously described.[15,34] Figure 1 shows the phylogenies of the HCV-1a and -1b (Figure 1*A*) and HCV-3a, -4a, -5a, and -6a (Figure 1*B*) sequences obtained in the present study along with closely related sequences. As shown in Figure 1*A*, 4 clusters of HCV-1b sequences were found, and some sequences from China, Korea, and the United Kingdom belonged to the Japanese *Sj* group. Also, 2 *Sj*-positive strains clustered with non-*Sj* strains in Japan. To measure country-wise clustering statistically, the isolation and migration of HCV-1b sequences were analyzed by use of a parsimony method. The estimated number of changes in location between groups (ie, migration events) was 7 for HCV-1b, whereas the expected number of location changes for the 1,000 simulated trees created with randomized locations was 40.38 for HCV-1b (Table 2). The observed number of migration events was significantly smaller ($P < .001$) than that expected under the null hypothesis of complete geographic mixing; therefore, this hypothesis can be rejected. This result suggests that there is considerable subdivision by location among the HCV-1b strains sampled. The parsimony analysis also provided clues about the movement of HCV-1b strains among the 6 groups. Table 2 shows the difference between the observed and expected number of changes for each pair of countries. In most cases, the observed number of migration events was

tapraid2/zg0-gast/zg0-gast/zg000306/zg02624d06z | xppws | S=1 | 1/27/06 | Art: 2624 | DTD5.0

Table 2. Isolation and Migration of HCV Subtype 1b Among 6 Groups

	Japan (*Sj*)	Korea	China	United Kingdom	Spain	Japan (non-*Sj*)
Number of observed changes in tree (total, 7)						
Japan (*Sj*)	-	1	2	1	0	0
Korea	0	-	0	0	0	0
China	0	0	-	0	0	0
United Kingdom	0	0	0	-	0	0
Spain	0	0	0	0	-	0
Japan (non-*Sj*)	2	0	0	0	0	-
Number of expected changes per tree (total, 40.38)						
Japan (*Sj*)	-	.09	3.30	.48	3.22	8.16
Korea	0	-	0	0	0	0
China	.02	0	-	.01	.63	1.36
United Kingdom	.02	0	.01	-	.01	.03
Spain	1.23	.02	.63	.09	-	1.39
Japan (non-*Sj*)	10.27	.13	4.26	.65	4.31	-
P value (total, $P < .001$)						
Japan (*Sj*)	-	NS	NS	NS	.027	< .001
Korea	NS	-	NS	NS	NS	NS
China	NS	NS	-	NS	NS	NS
United Kingdom	NS	NS	NS	-	NS	NS
Spain	NS	NS	NS	NS	-	NS
Japan (non-*Sj*)	< .001	NS	.012	NS	.011	-

smaller than expected, indicating no significant movement of HCV-1b strains among these groups. For HCV-1a and -3a, the observed number of migration events was significantly smaller than the expected number, again suggesting significantly less overall migration than would be expected by chance.

The Origin of HCV Subtypes in Each Country

Figure 2 shows a comparison of genetic distances estimated on all synonymous and nonsynonymous nucleotide positions between HCV strains in each country. The genetic distances were the greatest among HCV-1b strains in the Japanese *Sj*-positive group, followed by the Egyptian HCV-4a, Spanish HCV-1b, South African HCV-5a, and HCV-1b strains in the Japanese non-*Sj* group. The genetic distances among the US HCV-1a, the FSU HCV-3a, and Hong Kong's HCV-6a were similar and comparatively smaller. These data indicate that the Japanese HCV-1b population is the oldest, and Hong Kong's HCV-6a population is the youngest among the populations studied.

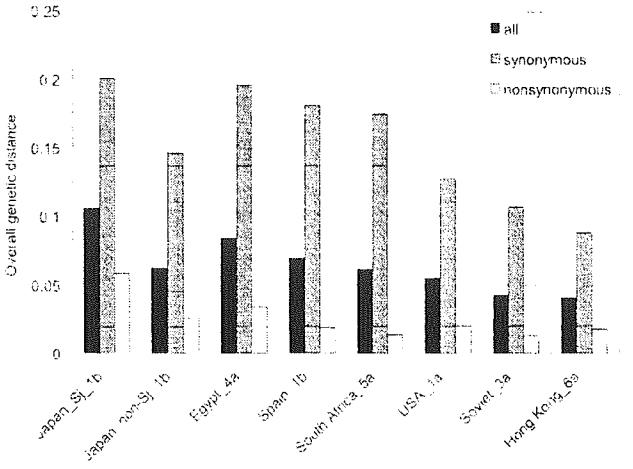

Figure 2. The overall mean genetic distances of all nucleotides positions, synonymous positions, and nonsynonymous positions in each country.

The molecular evolutionary rate was estimated by 2 independent methods. First, our previous linear regression analyses indicated that a molecular evolutionary rate was .58 (range, .55–.61) $\times 10^{-3}$ nucleotide substitutions/site/y.[11] Second, TipDate v1.2 was used to compare the different-rate model with the single rate and SRDT models. The SRDT model provides an adequate fit to most datasets ($P > .05$); the rates were similar (range, .50–.72) $\times 10^{-3}$ nucleotide substitutions/site/y. Although the SRDT model was rejected ($P < .05$) in 1 dataset, simulations have shown that even when the molecular clock is rejected, the confidence limits of the substitution rate sometimes may include the true rate.[26] Hence, we used the substitution rate estimated previously: .58 $\times 10^{-3}$ substitutions/site/y for the NS5B region.

Based on TipDate,[22] to investigate the origin of HCV subtypes we converted the genetic distance in the phylogenetic tree into a timescale by using the molecular clock. According to the timescale, the most recent com-

tapraid2/zg0-gast/zg0-gast/zg000306/zg02624d06z | xppws | S=1 | 1/27/06 | Art: 2624 | DTD5.0

mon ancestor for HCV subtypes in each country was established (Table 1); the divergence time of the most recent common ancestor for HCV-1b in the Japanese *Sj*-group was estimated before 1850, followed by that of Spanish HCV-1b strains (in 1892). The divergence time in the other groups was estimated to be in the 20th century. Notably, HCV-3a and HCV-6a strains have been introduced relatively recently into the FSU (in 1958) and Hong Kong (in 1963), respectively (Table 1).

Historical Analyses of the HCV Population by Using the Coalescent Theory

The level of population subdivision shown in the parsimony analysis described earlier suggests that much of the transmission of the sampled HCV strains occurred within the sampled groups. Therefore, the epidemic history of HCV strains in each group was estimated from separate trees.

Based on the phylogenetic analysis, the effective number of HCV infections through time, $N(t)$, was analyzed using a skyline plot for the HCV strains. The parameters for several models in Genie v3.5 were examined. Time (t) then was transformed to year using the same rate, assuming the collecting time to be the present. Figure 3 shows the skyline plots and HCV population growth in each country according to a piecewise expansion growth model that was evaluated by likelihood-ratio testing (data not shown).[24,25] An expansion growth model gave the best fit for only the *Sj*-positive HCV-1b population, but because this likelihood ratio was almost the same as that of the piecewise expansion growth model, all populations were applied to the same piecewise expansion growth model. Our estimates of the effective numbers of HCV infections showed a transition from constant size to rapid exponential growth in the 1920s among the Japanese *Sj*-positive HCV-1b population, as we have reported earlier.[15] This indicates the oldest outbreak among all studied populations, whereas the exponential growth among the Japanese *Sj*-negative HCV-1b population was dated in the 1940s,[15] which is close in time to the HCV-1b populations in Spain (Figure 3*A*) and other European countries,[27,28] and the HCV-4a population in Egypt.[18] The exponential growth of the HCV-5a population in South Africa occurred in the 1950s (Figure 3*B*), and comparatively recent HCV endemics were dated in the 1960s for both HCV-1a in the United States (Figure 3*C*) and HCV-3a in the FSU (Figure 3*D*), and in the late 1970s for HCV-6a in Hong Kong (Figure 3*E*).

The exponential growth rates also varied among the subtype populations (Table 1). The estimated rates for HCV-5a in South Africa, HCV-6a in Hong Kong, HCV-3a in the FSU, and HCV-1a in the United States were higher than those for HCV-1b in Japan, HCV-1b in Spain, and HCV-4a in Egypt. Hence, our findings indicate that the particular epidemics worldwide, associated with the corresponding HCV subtype, had different patterns in terms of divergence time, exponential spread time, and the dynamic growth rate. The different ages of the studied viral subpopulations, best assessed by synonymous genetic distance values, are shown in Figure 2.

Relative Population Size of HCV Subtypes in Each Country

Current estimates of the HCV subtype distribution in each sampled country were used to transform the epidemic histories shown in Figure 3 and previous data[1,4,15,18] to reflect the relative historical levels of HCV subtype infection in each country (Figure 4). As shown in Figure 4, 3 different growth patterns were found: one is the oldest historical pattern of HCV-1b in the Japanese *Sj* group, and the second group consists of HCV-1b in the Japanese non-*Sj* group, HCV-1b in Spain, and HCV-4a in Egypt. The last group, with the latest exponential growth, includes 4 different subtypes in independent countries: HCV-1a in the United States, HCV-3a in the FSU, HCV-5a in South Africa, and HCV-6a in Hong Kong.

Discussion

The world map of HCC occurrence still contains many wide gaps owing to difficulties collecting exact clinical and epidemiologic data from many countries. The positive correlation between HCV seroprevalence and HCC mortality was documented in a recent European report.[29] However, most of the HCV seroprevalence data worldwide are derived from studies of blood donors who represent a selective relatively low-risk, younger population. Therefore this approach underestimates the absolute burden of infection[30] and complicates a comparative analysis of the results obtained in different countries, especially for the age-specific HCV seroprevalence.

Because HCC is associated directly with the duration of HCV infection in a given carrier, the time of exposure to HCV infection is another critical determinant of HCC incidence at the population level. To investigate further the relative role played by the duration of HCV infection, we analyzed current and previously published data from populations throughout the world for the general and age-related HCV seroprevalence, the estimated time of HCV exponential spread, the association with primary risk factors for virus transmission, and HCC mortality.

The past population dynamics of a virus can be inferred from viral gene sequence data using a population