Research Article

Cancer Epidemiology, Biomarkers & Prevention

Distribution and Hepatocellular Carcinoma–Related Viral Properties of Hepatitis B Virus Genotypes in Mainland China: A Community-Based Study

Jianhua Yin[1], Hongwei Zhang[1], Yongchao He[1], Jiaxin Xie[1], Shijian Liu[1], Wenjun Chang[1], Xiaojie Tan[1], Chunying Gu[1], Wei Lu[1], Hongyang Wang[2], Shengli Bi[3], Fuqiang Cui[4], Xiaofeng Liang[4], Stephan Schaefer[5], and Guangwen Cao[1]

Abstract

Introduction: Hepatitis B virus (HBV) genotypes, replication status, and mutations have been associated with the risk of hepatocellular carcinoma (HCC). Our aim was to study the distribution and HCC-related viral properties of HBV genotypes/subgenotypes in Mainland China.

Methods: A multistage cluster probability sampling method was applied to select 81,775 participants between 1 and 59 years at 160 national disease surveillance points. We examined hepatitis B surface antigen, HBV genotypes and subgenotypes, hepatitis B e antigen, viral load, and mutations in the PreS and core promoter regions of HBV genome.

Results: HBV subgenotypes B2 (27.3%), C1 (10.7%), and C2 (58.0%) were predominant. Genotype D (D1, 80.8%) was frequent in the Uygur. We identified a new subgenotype, C9, mainly in Tibetans. Compositions of subgenotypes B2 and C1 and genotype mixture increased from the North to Central South, which was consistently associated with the increasing prevalence of hepatitis B surface antigen. Hepatitis B e antigen positivity and viral loads were higher in the young with genotype B and declined more rapidly with increasing age than those with genotype C. In contrast to G1896A, PreS deletion, T31C, T1753V, and A1762T/G1764A were more frequent in subgenotype C2 than in subgenotype B2. A1762T/G1764A, T1753V, C1653T, and G1896A, except PreS deletion, consecutively increased with increasing age.

Conclusion: HBV subgenotypes B2, C1, and C2 are endemic in Mainland China. HBV genotype C exhibits less replication activity in the young and harbors higher frequencies of the HCC-associated mutations than genotype B.

Impact: These basic data could help evaluate the association of HBV variations with HCC. *Cancer Epidemiol Biomarkers Prev; 19(3); 777–86. ©2010 AACR.*

Introduction

Hepatitis B virus (HBV) infection is a serious public health problem. Chronic HBV infection affects >300 million people worldwide. In Asia and most of Africa, chronic HBV infection is common and usually acquired perinatally or in childhood (1). Chronic HBV infection is by far the most important risk factor of hepatocellular carcinoma (HCC; ref. 2). Standard HBV vaccination dramatically decreases HCC prevalence among the vaccinees ages 6 to 19 years (3).

Eight genotypes (genotypes A-H) have been identified by a sequence divergence >8% in the entire HBV genome (4). HBV isolated in Vietnam and Laos has been suggested to form a ninth genotype I (5, 6). The designation has been questioned due to complex recombination (7). A HBV strain isolated from a Japanese patient has been provisionally designated HBV genotype J (8). Genotypes have further been separated into subgenotypes if the divergence in nucleotide sequence is between 4% and 8% (4). HBV genotypes have distinct geographic distributions and have been shown to differ with regard to clinical outcome, prognosis, and response to antiviral treatment (9-17). Infection with HBV genotype C has been frequently associated with an increased risk of HCC in the ages, whereas infection with genotype B has been found to be associated with the development of HCC and relapse of HCC in the young (9-14). Although

Authors' Affiliations: [1]Department of Epidemiology, [2]Laboratory for Signal Transduction, the 3rd Hospital, Second Military Medical University, Shanghai, China; [3]Departments of Hepatitis Virus, Institute for Viral Disease Control and Prevention, and [4]Immunization Programming, Chinese Center for Disease Control and Prevention, Beijing, China; and [5]Abteilung fur Virologie, Universitat Rostock, Germany

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

J. Yin and H. Zhang contributed equally to this work.

Corresponding Author: Guangwen Cao, Department of Epidemiology, Second Military Medical University, 800 Xiangyin Road, Shanghai 200433, People's Republic of China. Phone/Fax: 86-21-81871060. E-mail: gcao@smmu.edu.cn

doi: 10.1158/1055-9965.EPI-09-1001

HBV subgenotypes C1 and C2 are associated with an increased risk of HCC, only HBV subgenotype C2 is independently associated with HCC (9). Different HBV genotype/subgenotypes display distinct mutation patterns in the PreS (nucleotides 2848-154) region and in the enhancer II (EnhII, nucleotides 1636-1744)/basic core promoter (BCP, nucleotides 1751-1769)/precore region in the HBV genome. The precore/core encodes hepatitis B e surface antigen (HBeAg), which is used clinically as an indicator of active viral replication. HBeAg expression and high viral load are associated with an increased risk of HCC (18, 19). Some mutations in the PreS and the EnhII/BCP/Precore regions have been associated with the development of HCC (20-29). Our recent meta-analysis using published data up to August 31, 2008 has shown that HBV PreS mutations, C1653T, T1753V, and A1762T/G1764A, are associated with an increased risk of HCC (30). A1762T/G1764A is a valuable biomarker for identifying a subset of hepatitis B surface antigen (HBsAg) carriers who are at extremely high risk of HCC in prospective studies (22, 24, 27, 29).

In Mainland China, an endemic area with almost one third of the HBsAg carriers found worldwide (1), HBV genotypes and mutations have been investigated in a limited number of the hospital-based patients and selected geographic areas (21, 22, 24, 25, 27, 31, 32). However, no community-based epidemiologic study with sufficient participants has shown nationwide distribution of HBV genotypes and the prevalence of the HCC-associated viral mutations as well. In addition, the association between HBV genotype and viral replication status remains controversial because of conflicting data in literature. HBV genotype C is associated with increased viral load in Taiwan, whereas this genotype is associated with decreased viral load in Shanghai (13, 33). This difference is probably due to the age of the HBV-infected subjects as HBV DNA tends to fluctuate with time (11). An epidemiologic study with sufficient participants in a wide age range is necessary to address this issue. In this study, we conducted a large, community-based study to determine the distribution of HBV genotypes and its association with HCC-related viral properties in Mainland China. This study provides basic epidemiologic data of HBV genomic variations, which is useful for the evaluation of future HCC burden caused by chronic HBV infection and the development of strategy for the control of HCC.

Subjects and Methods

Study Population and Epidemiologic Survey

The study was carried out at 160 well-established National Disease Surveillance Points (DSP; a DSP was a study county) covering all 31 provincial administrative regions containing 22 provinces (Hebei, Shanxi, Liaoning, Jilin, Heilongjijang, Shaanxi, Gansu, Qinghai, Sichuan, Guizhou, Yunnan, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, and Hainan), 5 autonomous regions (Inner Mongolia, Guangxi, Ningxia, Xinjiang, and Tibet), and 4 municipalities directly under the central government (Beijing, Tianjin, Shanghai, and Chongqing) in Mainland China, from September to December in 2006. The representativeness of DSP is available at the Web site of DSP (34). The study subjects were randomly selected from community-based population at each DSP, regardless of their health states. To obtain enough study subjects for the evaluation of the distribution of HBV genotype and subgenotype in Mainland China, sample sizes at different age groups were estimated according to the estimated prevalence rates and allowable errors of HBsAg. The reason and procedure of sample size estimation are detailed in Supplementary Material. A multistage cluster probability sampling method was applied to select the study population. Briefly, two to four administrative townships were randomly selected from each DSP. One administrative village was randomly selected from each township. All children at the age between 1 and 4 y in the sampled villages were included in this study. Systematic sampling and family-based sampling methods were used to recruit the participants at the age between 5 and 14 y and those at the age between 15 and 59 y, respectively. A total of 81,775 participants, containing 16,376, 23,753, and 41,646 subjects at the ages between 1 and 4 y, between 5 and 14 y, and between 15 and 59 y, respectively, were involved in this study.

The participants were interviewed by trained research assistants using a standard questionnaire requesting information including sociodemographic characteristics. Fasting blood samples (2-4 mL) were collected with vacuum blood collection tubes (BD Diagnostics) without anticoagulant. The serum and blood clot were separated by centrifugation at 4°C at the Center for Disease Control and Prevention of each DSP, transported on dry ice and stored at −20°C in Chinese Center for Disease Control and Prevention, and at -40° C in the Department of Epidemiology, Second Military Medical University.

Informed consent in writing was obtained from each participant or guardian. Each resident who agreed to participate in this survey completed a questionnaire and provided blood sample. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Institutional Review Boards of Second Military Medical University and Chinese Center for Disease Control and Prevention.

Examination of HBV Serologic Markers and Serum Viral Load

All participants were examined for HBsAg. Those seropositive for HBsAg were examined for HBeAg. HBV serologic markers were examined by using ELISA (Xinchuang) according to the manufacturer's instructions. A total of 1,538 subjects (men, 837; women, 701)

with genotype determined were randomly selected for the examination of HBV viral load. Serum HBV DNA was measured in the LightCycler480 (Roche) using quantitative HBV PCR fluorescence diagnostic kits (Fosun Diagnostics). The kit has a certified lower limit of detection of 500 copies/mL. All measurements of serologic HBV markers including viral load were standardized by using corresponding Abbott reagents (Abbott laboratories). The results were sent to the participants within 2 to 3 wk after examination.

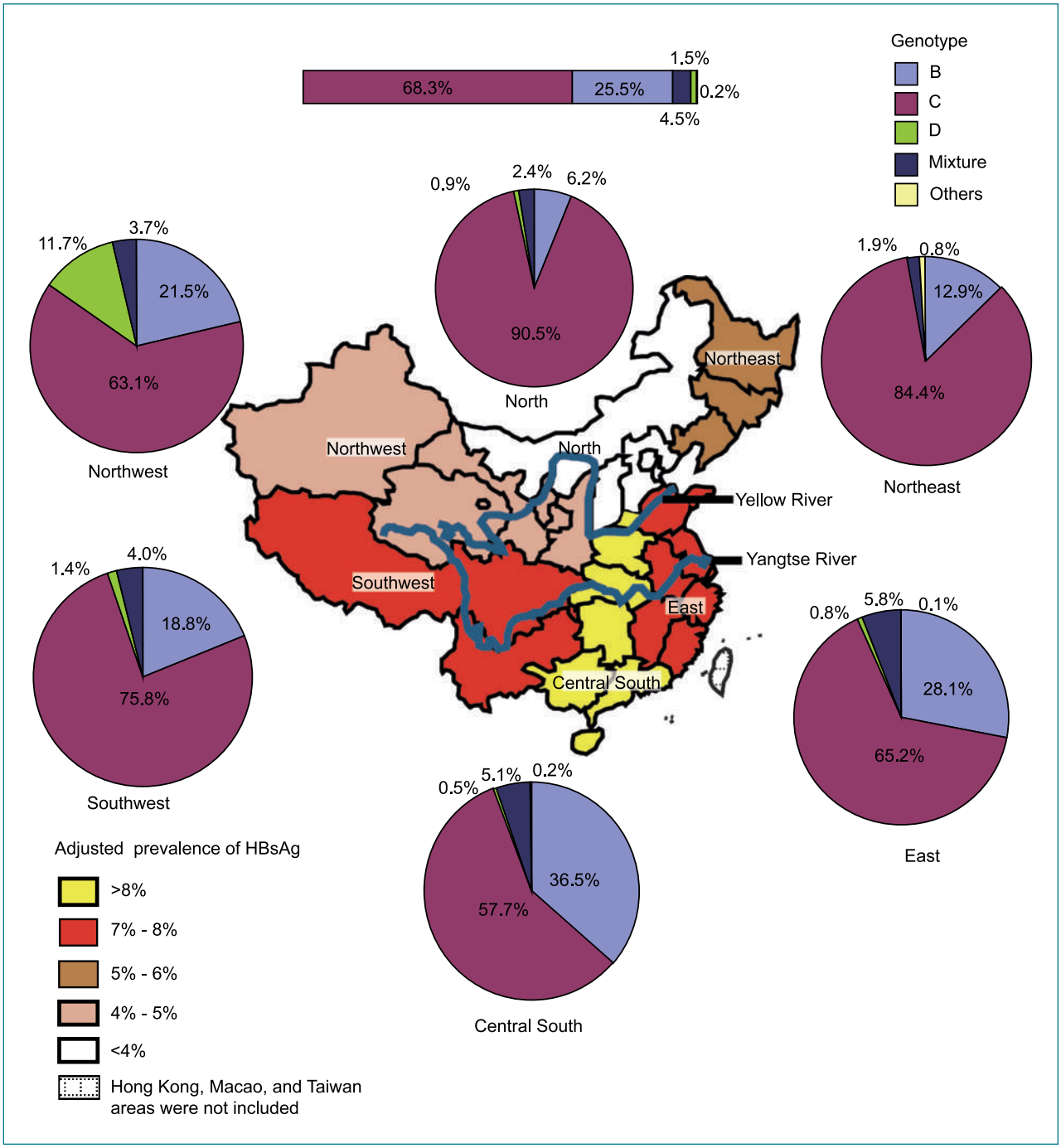

Figure 1. Geographic distribution of HBV genotypes and its association with HBsAg prevalence in six major regions of Mainland China. The North: Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia; the Northeast: Liaoning, Jilin, and Heilongjijang; the Northwest: Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang; the Southwest: Chongqing, Sichuan, Guizhou, Yunnan, and Tibet; the East: Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, and Shandong; Central South: Henan, Hubei, Hunan, Guangdong, Guangxi, and Hainan.

Table 1. The proportion of HBV genotypes B, C, and D and their subgenotypes in major topographical or geographic areas of Mainland China

Area	Genotype				Subgenotype												
	B*	C*	D	Total†	B1	B2*	B3	B4	C1*	C2*	C3	C4	C5	C9	D1	D2	Total
Area north to Yellow River	67 (11.3)	502 (84.5)	8 (1.3)	594 (100.0)	0	56 (11.5)	0	0	16 (3.3)	402 (82.5)	1 (0.2)	0	3 (0.6)	7 (1.4)	2 (0.4)	0	487 (100.0)
Area between Yangtse River and Yellow River	157 (26.0)	421 (69.8)	10 (1.7)	603 (100.0)	0	111 (25.7)	2 (0.5)	1 (0.2)	23 (5.3)	286 (66.2)	1 (0.2)	0	0	2 (0.5)	4 (0.9)	2 (0.5)	432 (100.0)
Area south to Yangtse River	495 (34.7)	833 (58.4)	9 (0.6)	1427 (100.0)	2 (0.2)	390 (40.0)	3 (0.3)	1 (0.1)	149 (15.3)	411 (42.1)	2 (0.2)	2 (0.2)	12 (1.2)	1 (0.1)	1 (0.1)	2 (0.2)	976 (100.0)
Uygur basin	6 (17.1)	10 (28.6)	17 (48.6)	35 (100.0)	0	5 (17.9)	0	0	0	8 (28.6)	0	0	0	0	14 (50.0)	1 (3.6)	28 (100.0)
Qinghai-Tibet plateau	8 (3.6)	204 (91.1)	2 (0.9)	224 (100.0)	0	6 (3.6)	0	0	29 (17.6)	109 (66.1)	0	0	0	21 (12.7)	0	0	165 (100.0)
Yunnan-Guizhou plateau	34 (26.2)	89 (68.5)	0	130 (100.0)	0	33 (28.9)	1 (0.9)	0	19 (16.7)	61 (53.5)	0	0	0	0	0	0	114 (100.0)
Total	767 (25.5)	2059 (68.3)	46 (1.5)	3013 (100.0)	2 (0.1)	601 (27.3)	6 (0.3)	2 (0.1)	236 (10.7)	1277 (58.0)	4 (0.2)	2 (0.1)	15 (0.9)	31 (1.4)	21 (1.0)	5 (0.2)	2202 (100.0)

NOTE: Data in the table were shown as *n* (%).

*P_{trend} < 0.001 from the Northern Yellow River area to the Southern Yangtse River area.

†The number includes HBV genotype mixture, genotypes A and E.

Determination of HBV Genotypes and Subgenotypes

HBV DNA was extracted from 200 μL HBsAg-positive sample using High Pure Viral Nucleic Acid kits (Roche Diagnostics) according to the manufacturer's instruction. Blood clots were also used as the sources of HBV DNA if insufficient serum was obtained as described in Supplementary Material. HBV genotypes and subgenotypes were determined by using a multiplex PCR assay (12, 35). HBV genotypes of samples with low levels of HBV DNA were identified by nested multiplex PCR as described in Supplementary Material. The primers used in this study are listed in Supplementary Table S1.

Sequencing, Mutation, Recombination, and Phylogenetic Analyses of HBV

Parts of the PreS and S gene region (PreS-S) of 24 novel HBV subgenotype isolates and the whole genomes of another 6 novel HBV subgenotype isolates were sequenced and deposited in Genbank. For viral mutation analysis, we sequenced the PreS region (from nt.2768 to nt.231) from 522 samples, and the EnhII/BCP/Precore region (from nt.1626 to nt.2004) from 750 samples of 1,000 subjects who were randomly selected from those infected with unique HBV subgenotype. These sequences were deposited in Genbank. Viral mutations were assessed by using the MEGA 4.0 software (36). Methods for PCR, DNA sequencing, and recombination analysis are detailed in Supplementary Material.

Statistical Analysis

χ^2 test was used to determine the differences in categorical variables, such as HBeAg positivities and the percentages of HBV genotypes. Continuous variables, such as serum viral loads with skewed distribution, were adjusted to normal distribution by transformation into logarithmic function and then tested by Student's *t* test or ANOVA. We also examined linear trends in the proportion of HBV genotypes or subgenotypes in different areas. All statistical tests were two sided and were performed by using the Statistical Program for Social Sciences (SPSS15.0 for Windows, SPSS). A *P* value of <0.05 was considered as statistically significant.

Results

Of 81,775 participants, 4,150 were HBsAg positive. A total of 1,371 of the 4,150 were HBeAg positive. The

prevalence rates (95% confidence intervals) of HBsAg expression at the age groups between 1 and 4 years, between 5 and 14 years, and between 15 and 59 years were 0.96% (0.75-1.17%), 2.42% (2.04-2.80%), and 8.57% (7.93-9.22%), respectively. Of the 4,150 HBsAg-positive samples, 379 had an inadequate preservation status. A total of 3,013 (79.90%) of the 3,771 HBsAg-positive samples were genotyped by using multiplex PCR either directly (n = 1,694) or as a second step of nested PCR (n = 1,319). No difference in the proportion of genotypes B and C was determined between the two methods (P = 0.13).

Geographic Distribution of HBV Genotypes and Subgenotypes

We found 767 subjects infected with HBV genotype B, 2,059 with genotype C, 46 with genotype D, 4 with genotype E, 1 with genotype A, and 136 with two or more genotypes (genotype mixture). HBV genotypes C (68.3%) and B (25.5%) were predominant. The distribution of HBV genotypes showed obvious geographic characteristics (Fig. 1). Genotype D was endemic in the Northwest. The proportion of genotype B consecutively increased, whereas genotype C decreased, from North to Yellow River to South to Yangtse River (Table 1). The proportions of genotype B in North versus South to Yangtse River were 18.3% versus 30.65%, in contrast to genotype C (75.9% versus 62.9%, $P < 0.001$). The proportions of genotype B and genotype mixture consecutively increased across the North, the Northeast, the Northwest, the Southwest, the East, and Central South, and were consistently associated with increasing prevalence of HBsAg, in contrast to genotype C ($P_{trend} < 0.001$; Supplementary Table S2; Fig. 1).

HBV subgenotype B2 (98.4%) was the major subgenotype of HBV genotype B. HBV subgenotype C2 (80.1%)

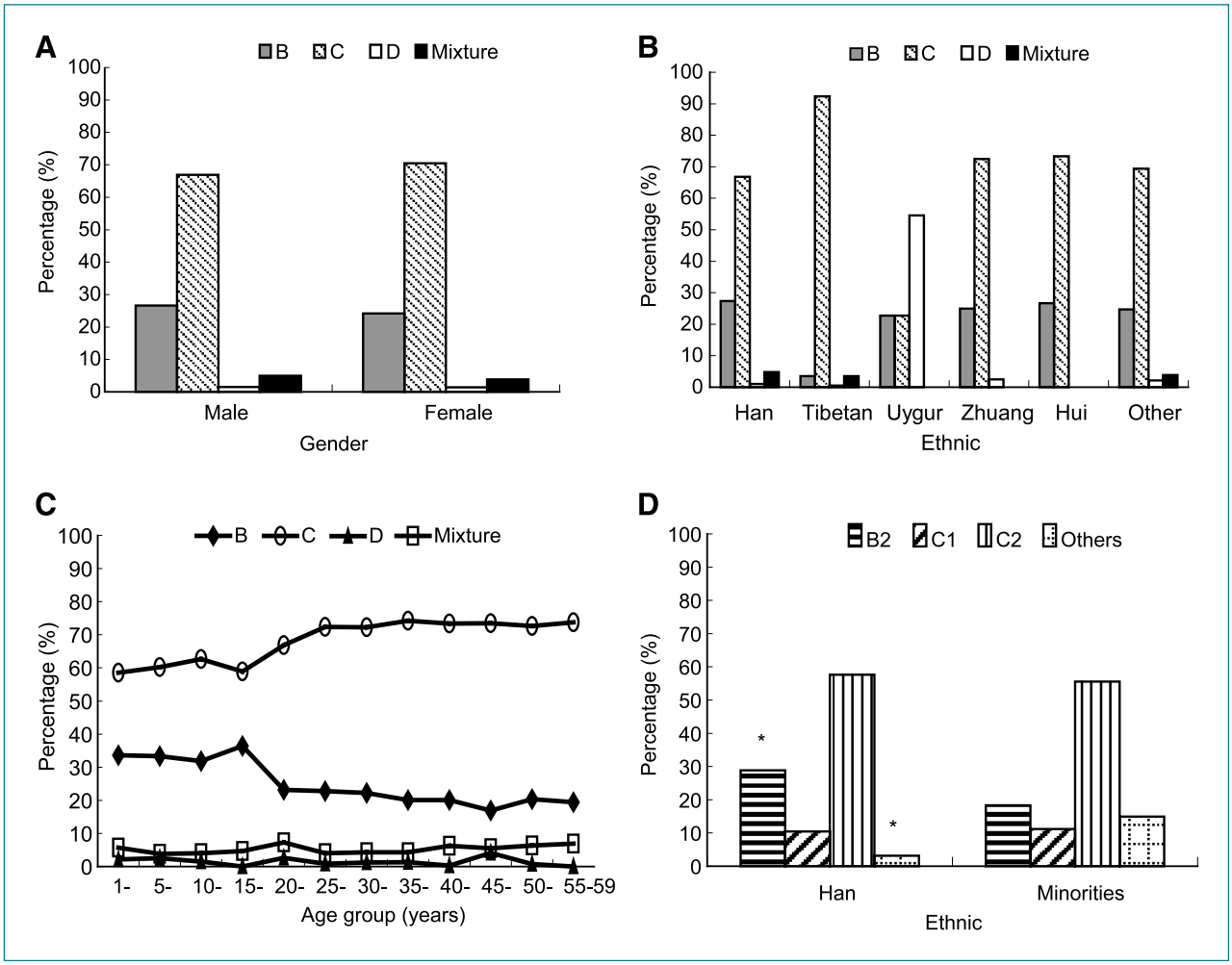

Figure 2. Demographic distribution of HBV genotypes in Mainland China. A, the proportion of HBV genotypes in men and women; B, the proportion of HBV genotypes in the major ethnicities; C, the proportion of HBV genotypes at each age group. D, the proportion of HBV subgenotypes in the Han and the minorities in combination. *, $P < 0.001$ between the Han and the minorities in combination.

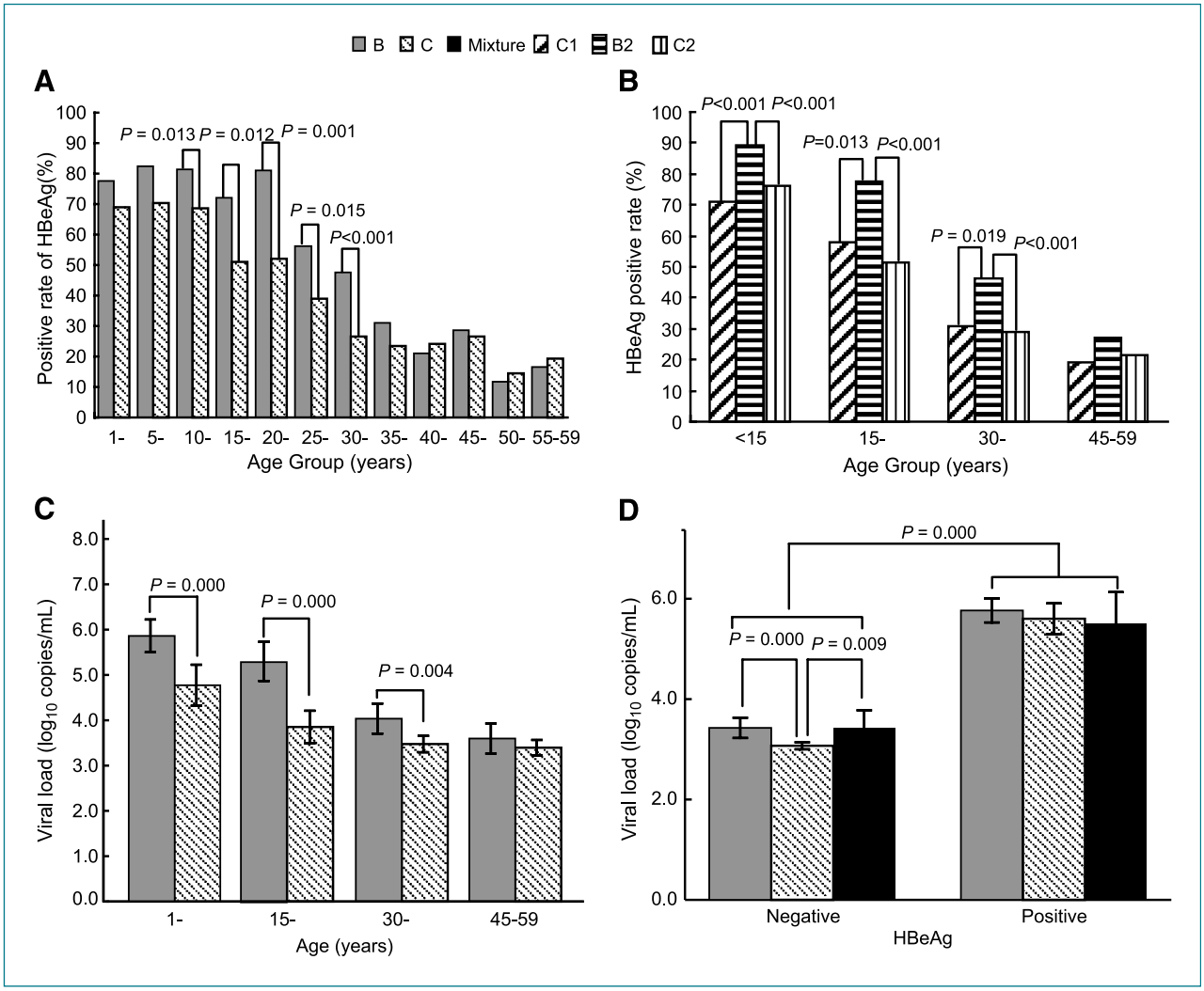

Figure 3. HBeAg seropositivities and serum viral loads of the subjects infected with HBV genotype B or genotype C. A, HBeAg positivities of the subjects with genotype B and those with genotype C at each age group; B, HBeAg seropositivities of the subjects with subgenotype B2, the subjects with subgenotype C1, and those with subgenotype C2 at each age group; C, serum viral loads of the subjects with HBV genotype B and those with genotype C at each age group; D, serum viral loads of the subjects with genotype B and those with genotype C at different HBeAg status. B, genotype B; C, genotype C; Mix, genotype mixture.

and HBV subgenotype C1 (14.8%) were the major subgenotypes of genotype C. The proportions of subgenotypes C1 and B2 consecutively increased from the North to Yellow River to the South to Yangtse River and consecutively increased from the North to Central South, in contrast to subgenotype C2 (Supplementary Table S3; Table 1). We identified a new subgenotype of genotype C, mostly found in Qinghai-Tibet plateau (Table 1). The whole genomes from six samples, PreS-S fragments from another 24 samples, and the EnhII/BCP/Precore fragment from one more sample were sequenced. The divergence in the whole genome was from 4.1% to 6.7% between the novel subgenotype and each of subgenotypes C1 to C5, and over 8% between the subgenotype and each of the other genotypes (Supplementary Table S4). The phylogenetic analysis with the whole sequences of HBV showed that it represented a new clade of genotype C (Supplementary Fig. S1). The SimPlot analysis indicated that it did not show gross recombination but was made up of smaller fragments sharing >50% homology with HBV subgenotypes D2, D4, C1, C2, C4, and C6 (in correct numerical order C8; Supplementary Fig. 2; ref. 37). HBV C7 and C6 were recently designated in Indonesia (38, 39). This subgenotype was then designated as HBV subgenotype C9. We found 2 subjects infected with subgenotype B1, 6 with B3, 2 with B4, 4 with C3, 2 with C4, and 15 with C5, mostly from Southern China. We also found 21 subjects infected with subgenotype D1 and 5 with D2, mostly from the Northwest (Supplementary Table S5).

Demographic Distribution of HBV Genotypes and Subgenotypes

The distribution of HBV genotypes was not significantly different between men and women (Fig. 2A). More genotype D (54.6%) and less genotype C (22.7%) were found in the Uygur than in other ethnicities ($P < 0.001$; Fig. 2B). Genotype mixture was rare in the minorities. Genotype B was more prevalent in the young than in those at the age of 25 years or older, in contrast to genotype C ($P < 0.001$; Fig. 2C). In contrast to subgenotypes B2 and C1, HBV subgenotype C2 was more prevalent in the Mandarin-speaking population than in the Cantonese- or other Southern dialect–speaking population ($P < 0.001$). Of the 31 subjects with subgenotype C9, 21 (67.74%) were Tibetan. In contrast to subgenotype B2, the proportions of subgenotypes C4 ($P = 0.030$), C5 ($P < 0.001$), C9 ($P < 0.001$), and D1 ($P < 0.001$) were significantly higher in the minorities in combination than in the Han. Figure 2D shows the difference in the proportion of HBV subgenotypes between the Han and the minorities. Subgenotype D1 was mainly found in the Uygur (Supplementary Table S6). PreS-S fragment of subgenotype D1 from the Uygur showed a high similarity to strains from Turkey by phylogenetic analysis (Supplementary Fig. S3).

HBeAg Status and Serum Viral Loads of the Subjects Infected with Different HBV Genotypes and Subgenotypes

Overall, the genotype B HBV–infected subjects were significantly more often HBeAg positive than the genotype C HBV–infected subjects ($P < 0.001$), regardless of the ethnicity and gender. The HBeAg-negative subjects were significantly older than the HBeAg-positive subjects ($P < 0.001$). The subjects with genotype B were significantly more often HBeAg positive than those with genotype C at the age between 10 and 35 years; however, no significant difference in HBeAg positivity between the subjects with genotype B and those with genotype C was found in those older than 35 years (Fig. 3A). The subjects with HBV subgenotype B2 were significantly more often HBeAg positive than the subjects with subgenotype C1 or those with subgenotype C2 at the age between 1 and 45 years, whereas no difference in HBeAg positivity between the subjects with HBV subgenotype B2 and the subjects with subgenotype C1 or C2 was found in those older than 45 years (Fig. 3B). Serum viral load was significantly higher in the subjects with genotype B than in those with genotype C at the age younger than 45 years, whereas this difference in viral load was not found in those older than 45 years (Fig. 3C). Serum HBV DNA was significantly higher in

Table 2. Frequencies of the HCC and other liver disease–associated HBV mutations in the randomly selected subjects infected with unique HBV subgenotype

Genotype	PreS				EnhII		BCP and Precore					
	n	Deletion*	T31C†	T53C‡	*n*	C1653T	*n*	T1753V§	A1762T/ G1764A‖	C1766T	G1896 A¶	G1899 A**
B	151	11 (7.3)	4 (2.7)	9 (6.0)	143	2 (1.4)	214	0	15 (7.0)	0	17 (8.0)	10 (4.7)
B2	145	10 (6.9)	4 (2.8)	8 (5.5)	135	2 (1.5)	206	0	14 (6.8)	0	13 (6.3)	9 (4.4)
C	353	47 (13.3)	76 (21.5)	4 (1.1)	435	15 (3.5)	514	28 (5.5)	76 (14.8)	11 (2.1)	16 (3.1)	4 (0.8)
C1	47	5 (10.6)	1 (2.1)	0	66	1 (1.5)	72	6 (8.3)	17 (23.6)	1 (1.4)	2 (2.8)	1 (1.4)
C2	268	38 (14.2)	75 (28.0)	4 (1.5)	328	13 (4.0)	395	18 (4.6)	51 (12.9)	9 (2.3)	13 (3.3)	3 (0.8)
C5	9	2 (22.2)	0	0	13	0	13	1 (7.7)	2 (15.4)	0	1 (7.7)	0
C7	24	2 (8.3)	0	0	23	1 (4.4)	29	3 (10.3)	5 (17.2)	1 (3.4)	0	0
D	18	13 (72.2)	0	0	22	0	22	0	1 (4.6)	0	2 (9.1)	0
D1	14	11 (78.6)	0	0	18	0	18	0	1 (5.6)	0	2 (11.1)	0
Total	522	71 (13.6)	80 (15.3)	13 (2.5)	600	17 (2.8)	750	28 (3.7)	92 (12.3)	11 (1.5)	35 (4.7)	14 (1.9)

NOTE: Data in the table were shown as *n* (%). We identified six participants infected with HBV B3, two with HBV B4, three with HBV C3, two with HBV C4, and four with HBV D2. Data of HBV mutation from these participants were not listed in the table because of small size.

*PreS deletion, B vs C $P = 0.05$; B2 vs C2 $P = 0.03$; $P < 0.001$ for B vs D, C vs D, D1 vs B2, D1 vs C1, D1 vs C2, and D1 vs C7; D1 vs C5 $P = 0.008$.

†T31C, $P < 0.010$ for B vs C, B2 vs C2, C2 vs C1, and C2 vs C7; C2 vs D1 $P = 0.020$.

‡T53C, B vs C $P = 0.005$; B2 vs C2 $P = 0.02$.

§T1753V, $P < 0.001$ for B vs C, B2 vs C1, B2 vs C2, B2 vs C7, and B2 vs C5.

‖A1762T/G1764A, $P < 0.01$ for B vs C and B2 vs C1; $P < 0.05$ for B2 vs C2 and C1 vs C2.

¶G1896A, B vs C $P = 0.004$.

**G1899 A, B vs C $P = 0.001$; B2 vs C2 $P = 0.003$.

Table 3. Frequencies of the HCC and other liver disease–associated HBV mutations in the randomly selected subjects with increasing age

Age (y)	PreS			EnhII	BCP and precore				
	Deletion	T31C	T53C*	C1653T†	T1753V‡	A1762T/G1764A‡	C1766T§	G1896A‡	G1899A
<15	16.1 (31/193)	8.8 (17/194)	0.5 (1/194)	1.3 (3/235)	1.5 (4/265)	6.0 (16/265)	0.8 (2/265)	1.1 (3/265)	0 (0/265)
15-29	13.3 (17/128)	19.4 (28/129)	0.8 (1/129)	2.2 (3/139)	1.9 (3/162)	11.7 (19/162)	0.6 (1/162)	3.1 (5/162)	2.5 (4/162)
30-44	10.7 (15/140)	15.7 (22/140)	5.0 (7/140)	2.6 (4/152)	5.2 (11/210)	12.9 (27/210)	1.9 (4/210)	8.1 (17/210)	1.0 (2/210)
45-59	14.5 (10/69)	18.8 (13/69)	5.8 (4/69)	9.1 (7/77)	8.5 (10/118)	25.4 (30/118)	3.4 (4/118)	10.2 (12/118)	6.8 (8/118)

NOTE: Data were shown as % (*n*).
*P_{trend} = 0.002.
†P_{trend} = 0.003.
‡P_{trend} < 0.001.
§P_{trend} = 0.041.

the HBeAg-positive subjects than in the HBeAg-negative subjects. Furthermore, serum HBV DNA was significantly higher in the HBeAg-negative subjects with genotype B or the HBeAg-negative subjects with genotype mixture than in those with genotype C (Fig. 3D).

Liver Disease–Associated Mutation Patterns of Each HBV Genotype and Subgenotype from HBV-Infected Subjects

After sequencing the PreS region from 522 samples and the EnhII/BCP/Precore region from 750 samples, we rated the HCC-associated HBV mutations including PreS deletion, T31C, T53C, C1653T, T1753V, A1762T/G1764A, and G1899 A (20-31); fulminant hepatitis B–associated mutation G1896A (40); and liver cirrhosis–associated mutations C1766T (20). The mutation rates of HBV genotypes and subgenotypes are summarized in Table 2. T1753V, A1762T/G1764A, and T31C were more frequent in genotype C than in genotype B, in contrast to T53C, G1896A, and G1899 A. The frequencies of T31C, and A1762T/G1764A were significantly different between HBV subgenotype C1 and subgenotype C2. PreS deletion, T31C, T1753V, and A1762T/G1764A were more frequent in HBV subgenotype C2 than in subgenotype B2. Surprisingly, 27.8% of genotype D isolates did not harbor a 33-bp deletion in the PreS1 region, a hallmark of all known genotype D (4).

We evaluated the combination of HBV mutations including PreS deletion, C1653T, T1753V, A1762T/G1764, and G1896A. No significant differences in the frequencies of the combined mutations were found between genotype B and genotype C (data not shown). We then stratified the subjects with HBV mutation information into age groups. The frequencies of A1762T/G1764A, T1753V, C1653T, T53C, C1766T, and G1896A increased consecutively with increasing age, whereas the frequencies of PreS deletion, T31C, and G1899A did not increase significantly with increasing age (Table 3). The frequency of PreS deletion, rather than other mutations, was significantly higher in male than female (P = 0.010). The frequencies of C1653T, T1753V, A1762T/G1764A, and G1896A increased consecutively with increasing age in the genotype C HBV–infected group (Supplementary Table S7).

Discussion

In this study, we recruited a large study population within the 160 DSPs, which were demographically and geographically representative for Mainland China. We used sensitive techniques for genotyping. The ratio of HBV genotypes identified by the nested multiplex-PCR was similar to that identified by routine multiplex PCR. Thus, the distribution of HBV genotypes from this study is also representative for HBV carriers with low viremia and does not have bias even if levels of viremia would be different among genotypes or age groups.

Genotypes C and B are the major HBV genotypes endemic in Mainland China. However, their distribution is geographically and demographically different. In contrast to HBV subgenotype C2, HBV subgenotypes B2, C1, and genotype mixture increase from the North to Central South. These results indicate that HBV subgenotype C2 has probably evolved in North China, whereas subgenotypes B2 and C1 have probably evolved in South China. Subgenotypes B3, C3, C4, and C5 found in South China might have been introduced from South Asia where these strains are endemic (41). Genotype D inherently harbors a 33-bp deletion in the PreS1 region. We found that 27.8% of HBV genotype D isolated in Northwest China did not harbor this deletion (Table 2), which might represent an ancestral form of genotype D that has been preserved for unknown reasons. Subgenotype D1 is endemic in the Mediterranean region including Turkey (42). Subgenotype D1 isolated from the Uygur was phylogenetically linked to the strains in Turkey. Thus, we speculate that subgenotype D1 has probably evolved in Central Asia where the Uygur lived and has spread to

Turkey, Iran, and entire Northern Eurasia. We identified a new subgenotype C9 that differed from recombinant genomes made up of parts of genotypes C and D described in Tibet (43). Presumably, C9 has evolved in Tibet because 67.74% of the C9 HBV-infected subjects were Tibetan.

HBeAg seropositivity and serum viral load were high in the young with the genotype B (mainly B2), indicating that genotype B replicates more actively in the young. High replication rates of genotype B in young adults lead to an increased risk of horizontal transmission of HBV by sexual activity because high concentration of HBV DNA in serum is associated with high concentrations in semen and other body fluids of HBV carriers (44). Genotype B has recently been shown by us to be more apt to cause acute hepatitis B (33). In contrast to genotype C, the proportion of genotype B was significantly higher in those younger than 25 years than in the older age group (Fig. 2C). HBeAg seropositivity and serum viral load decreased more rapidly in the genotype B HBV–infected subjects than in the genotype C HBV–infected subjects with increasing age (Fig. 3A-C). Genotype B is associated with earlier HBeAg seroconversion than genotype C (45, 46). Early HBeAg seroconversion typically confers a favorable outcome (47). These evidence indicate that genotype B is more apt to be earlier cleared than genotype C, although further evidence are needed to address this issue.

This study showed that different HBV genotype/subgenotypes had different patterns of HCC-associated HBV mutations in the PreS and EnhII/BCP/Precore regions. PreS deletion, T31C, T1753V, and A1762T/G1764A were more frequent in subgenotype C2 than in subgenotype B2 (Table 2). However, the frequencies of the combined mutations were not significantly different between genotype C and genotype B. PreS deletion, T53C, T31C, C1653T, T1753V, and A1762T/G1764A have been associated with an increased risk of HCC. We have shown that the frequencies of PreS deletion, C1653T, T1753V, and A1762T/G1764A accumulate during the progression of chronic HBV infection from the asymptomatic carrier state to liver cirrhosis or HCC (30). Although genotype B is associated with HCC or HCC recurrence in young, mostly noncirrhotic, patients (10-12), infection with HBV genotype C is associated with the increased risks of cirrhosis and HCC at older age compared with infection with HBV genotype B (13, 14). This difference in the relevance of HBV genotypes B and C with the development of liver cirrhosis and/or HCC might be due to the distinct mutation patterns between genotypes. This study also revealed that G1896A was more frequent in genotype B than in genotype C (Table 2). G1896A has been linked to a decreased risk of HCC with increasing age in a prospective study (29). The frequencies of A1762T/G1764A, T1753V, C1653T, T53C, and G1896A increased consecutively with increasing age (Table 3). Thus, HBV subgenotype C2 might be more apt to cause liver cirrhosis and HCC than subgenotype B2 with increasing age. The high percentage of genotype C in Mainland China indicates the challenge of liver cirrhosis and HCC caused by chronic HBV infection in the coming decades.

Several limitations should be addressed. The design of this study is cross-sectional in nature. Thus, the HBV detected could be a newly acquired or a persistent infection. Also, we were unable to perform liver function tests and the examination of liver disease status for all participants at the time of field survey, resulting in loss of data.

In summary, this study first provides basic epidemiologic data on demographic and geographic distributions and the HCC-associated viral properties of HBV genotypes and subgenotypes in community-based populations in Mainland China, the largest HBV endemic area with 1.3 billion people and 56 ethnicities. The distribution and the viral properties of HBV genotypes endemic in Mainland China are internationally important because of immigration, cheap air travel, and globalization (48). HBV genotypes and the HCC-associated mutations in asymptomatic HBsAg carriers are of high epidemiologic significance because they are unrecognized sources of infection and will be the future patients with HCC. Early interventions to the HBV-infected people with high risk should be helpful in decreasing overall incidence of HCC.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Prof. Wolfram H. Gerlich (Justus-Liebig University of Giessen, Germany) for his critical review of this work, and Junwen Sun, Liping Lin, Jinfeng Zhao, Jinsong Chen, Lei Han, and Wenying Lu (Department of Epidemiology, Second Military Medical University) for their help with laboratory work.

Grant Support

Ministry of Health of China grant 2004BA718B01 (G. Cao), 2008ZX10002-15 (G. Cao), National Natural Science Foundation of China grant 30921006 (G. Cao), the Shanghai Science & Technology Committee grant 08XD14001 (G. Cao), and the Shanghai Board of Health grants 08GWD02 (G. Cao), 08GWZX0201, (G. Cao), and Deutsche Forschungsgemeinschaft grant Scha778/3-1 (S. Schaefer).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 09/24/2009; revised 11/16/2009; accepted 12/31/2009; published OnlineFirst 02/16/2010.

References

1. Lai CL, Ratziu V, Yuen MF, Poynard T. Viral hepatitis B. Lancet 2003; 362:2089–94.

2. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 2007;132:2557–76.

3. Chang MH, You SL, Chen CJ, et al. Decreased incidence of hepatocellular carcinoma in hepatitis B vaccinees: a 20-year follow-up study. J Natl Cancer Inst 2009;101:1348–55.
4. Schaefer S. Hepatitis B virus: significance of genotypes. J Viral Hepat 2005;12:111–24.
5. Tran TT, Trinh TN, Abe K. New complex recombinant genotype of hepatitis B virus identified in Vietnam. J Virol 2008;82:5657–63.
6. Olinger CM, Jutavijittum P, Hubschen JM, et al. Possible new hepatitis B virus genotype, southeast Asia. Emerg Infect Dis 2008;14: 1777–80.
7. Kurbanov F, Tanaka Y, Kramvis A, Simmonds P, Mizokami M. When should "I" consider a new hepatitis B virus genotype? J Virol 2008;82: 8241–2.
8. Tatematsu K, Tanaka Y, Kurbanov F, et al. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. J Virol 2009;83:10538–47.
9. Chan HL, Tse CH, Mo F, et al. High viral load and hepatitis B virus subgenotype ce are associated with increased risk of hepatocellular carcinoma. J Clin Oncol 2008;26:177–82.
10. Ni YH, Chang MH, Wang KJ, et al. Clinical relevance of hepatitis B virus genotype in children with chronic infection and hepatocellular carcinoma. Gastroenterology 2004;127:1733–8.
11. Yin J, Zhang H, Li C, et al. Role of hepatitis B virus genotype mixture, subgenotypes C2 and B2 on hepatocellular carcinoma: compared with chronic hepatitis B and asymptomatic carrier state in the same area. Carcinogenesis 2008;29:1685–91.
12. Kao JH, Chen PJ, Lai MY, Chen DS. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. Gastroenterology 2000;118:554–9.
13. Yu MW, Yeh SH, Chen PJ, et al. Hepatitis B virus genotype and DNA level and hepatocellular carcinoma: a prospective study in men. J Natl Cancer Inst 2005;97:265–72.
14. 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.
15. Erhardt A, Blondin D, Hauck K, et al. Response to interferon alfa is hepatitis B virus genotype dependent: genotype A is more sensitive to interferon than genotype D. Gut 2005;54:1009–13.
16. Wai CT, Chu CJ, Hussain M, Lok AS. HBV genotype B is associated with better response to interferon therapy in HBeAg(+) chronic hepatitis than genotype C. Hepatology 2002;36:1425–30.
17. Erhardt A, Gobel T, Ludwig A, et al. Response to antiviral treatment in patients infected with hepatitis B virus genotypes E-H. J Med Virol 2009;81:1716–20.
18. Chen CJ, Yang HI, Su J, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. JAMA 2006;295:65–73.
19. Yang HI, Lu SN, Liaw YF, et al. Hepatitis B e antigen and the risk of hepatocellular carcinoma. N Engl J Med 2002;347:168–74.
20. Chen CH, Hung CH, Lee CM, et al. Pre-S deletion and complex mutations of hepatitis B virus related to advanced liver disease in HBeAg-negative patients. Gastroenterology 2007;133:1466–74.
21. Gao ZY, Li T, Wang J, et al. Mutations in preS genes of genotype C hepatitis B virus in patients with chronic hepatitis B and hepatocellular carcinoma. J Gastroenterol 2007;42:761–8.
22. Yuan JM, Ambinder A, Fan Y, Gao YT, Yu MC, Groopman JD. Prospective evaluation of hepatitis B 1762(T)/1764(A) mutations on hepatocellular carcinoma development in Shanghai, China. Cancer Epidemiol Biomarkers Prev 2009;18:590–4.
23. Chou YC, Yu MW, Wu CF, et al. Temporal relationship between hepatitis B virus enhancer II/basal core promoter sequence variation and risk of hepatocellular carcinoma. Gut 2008;57:91–7.
24. Chen JG, Kuang SY, Egner PA, et al. Acceleration to death from liver cancer in people with hepatitis B viral mutations detected in plasma by mass spectrometry. Cancer Epidemiol Biomarkers Prev 2007;16: 1213–8.
25. Sung JJ, Tsui SK, Tse CH, et al. Genotype-specific genomic markers associated with primary hepatomas, based on complete genomic sequencing of hepatitis B virus. J Virol 2008;82:3604–11.
26. Yuen MF, Tanaka Y, Shinkai N, et al. Risk for hepatocellular carcinoma with respect to hepatitis B virus genotypes B/C, specific mutations of enhancer II/core promoter/precore regions and HBV DNA levels. Gut 2008;57:98–102.
27. Fang ZL, Sabin CA, Dong BQ, et al. HBV A1762T, G1764A mutations are a valuable biomarker for identifying a subset of male HBsAg carriers at extremely high risk of hepatocellular carcinoma: a prospective study. Am J Gastroenterol 2008;103:2254–62.
28. Kao JH, Chen PJ, Lai MY, Chen DS. Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. Gastroenterology 2003;124:327–34.
29. Yang HI, Yeh SH, Chen PJ, et al. Associations between hepatitis B virus genotype and mutants and the risk of hepatocellular carcinoma. J Natl Cancer Inst 2008;100:1134–43.
30. Liu S, Zhang H, Gu C, et al. Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. J Natl Cancer Inst 2009;101:1066–82.
31. Yuan J, Zhou B, Tanaka Y, et al. Hepatitis B virus (HBV) genotypes/subgenotypes in China: mutations in core promoter and precore/core and their clinical implications. J Clin Virol 2007;39:87–93.
32. Zeng G, Wang Z, Wen S, et al. Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotypes in China. J Viral Hepat 2005;12:609–17.
33. Zhang H, Yin J, Li Y, et al. Risk factors and chronicification of acute hepatitis caused by hepatitis B virus genotypes B2 and C2 in Shanghai, China. Gut 2008;57:1713–20.
34. National Disease Surveillance Points (DSPs) of Mainland China. Available from: http://www.chinadsps.com/yufang/en_jianjie1.asp
35. Chen J, Yin J, Tan X, et al. Improved multiplex-PCR to identify hepatitis B virus genotypes A-F and subgenotypes B1, B2, C1 and C2. J Clin Virol 2007;38:238–43.
36. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24:1596–9.
37. Cavinta L, Cao G, Schaefer S. Description of a new hepatitis B virus C6 subgenotype found in the Papua province of Indonesia and suggested renaming of a tentative C6 subgenotype found in the Philippines as subgenotype C7. J Clin Microbiol 2009;47:3068–9.
38. Lusida MI, Nugrahaputra VE, Soetjipto , et al. Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. J Clin Microbiol 2008;46:2160–6.
39. Mulyanto , Depamede SN, Surayah K, et al. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. Arch Virol 2009;154: 1047–59.
40. Ozasa A, Tanaka Y, Orito E, et al. Influence of genotypes and precore mutations on fulminant or chronic outcome of acute hepatitis B virus infection. Hepatology 2006;44:326–34.
41. 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.
42. Bozdayi G, Turkyilmaz AR, Idilman R, et al. Complete genome sequence and phylogenetic analysis of hepatitis B virus isolated from Turkish patients with chronic HBV infection. J Med Virol 2005;76: 476–81.
43. Cui C, Shi J, Hui L, et al. The dominant hepatitis B virus genotype identified in Tibet is a C/D hybrid. J Gen Virol 2002;83:2773–7.
44. Kidd-Ljunggren K, Holmberg A, Blackberg J, Lindqvist B. High levels of hepatitis B virus DNA in body fluids from chronic carriers. J Hosp Infect 2006;64:352–7.
45. Livingston SE, Simonetti JP, Bulkow LR, et al. Clearance of hepatitis B e antigen in patients with chronic hepatitis B and genotypes A, B, C, D, and F. Gastroenterology 2007;133:1452–7.
46. Chu CJ, Hussain M, Lok AS. Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. Gastroenterology 2002;122:1756–62.
47. Lin CL, Kao JH. Hepatitis B viral factors and clinical outcomes of chronic hepatitis B. J Biomed Sci 2008;15:137–45.
48. Williams R. Global challenges in liver disease. Hepatology 2006;44: 521–26.