Hepatitis B virus subgenotype A1 predominates in liver disease patients from Kerala, India

Deepak Gopalakrishnan, Mark Keyter, Kotacherry Trivikrama Shenoy, Kondarappassery Balakumaran Leena, Lakshmikanthan Thayumanavan, Varghese Thomas, KR Vinayakumar, Charles Panackel, Arun T Korah, Ramesh Nair, Anna Kramvis

Abstract

AIM: To molecularly characterize hepatitis B virus (HBV) isolates from Kerala and to relate them to the clinical manifestation of infection.

METHODS: Sera and clinical data were collected from 91 patients diagnosed with chronic HBV infection and HBV-related hepatocellular carcinoma (HCC). HBV from 44 HCC, 22 cirrhotic and 25 chronic hepatitis patients were genotyped by sequencing of the complete S region or by restriction fragment length polymorphism assays. The basic core promoter/precore region was sequenced. The complete surface DNA sequences were assembled and aligned manually, and then compared with the sequences of HBV of genotypes (A-J) from GenBank. The evolutionary history was inferred using the Neighbor-Joining method and the evolutionary distances computed using the Kimura 2-parameter method. Bootstrapping was performed using 1000 replicates. The TaqMan BS-1 probe was used to quantify HBV DNA at a lower detection limit of approximately 20 IU/mL. Continuous variables were compared using an independent Student’s t test. The χ² test or Fisher’s exact test was used to compare categorical variables. The differences were considered statistically significant at P < 0.05.

RESULTS: Irrespective of disease status, the predominant genotype was A (72%); 95% belonging to subgenotype A1, followed by genotypes D (27%) and C (1%). HCC patients infected with subgenotype A1 were significantly younger than those infected with D. Mutation A1762T/G1764A was significantly associated with HCC in both genotypes A and D. Mutation G1862T was more frequent in subgenotype A1 (P < 0.0001), and in combination with A1762T/G1764A, it was significantly associated with HBV from HCC patients. Mutation C1766T/T1768A was significantly associated with...
genotype A \((P = 0.05)\) and HCC \((P = 0.03)\). The preS2 start codon M1T/I mutation was unique to genotype A strains (15.6\%) from all disease groups and occurred at a higher frequency in isolates from HCC patients \((P = 0.076)\). A higher frequency of preS deletion mutants (33.3\%) was observed in genotype A from HCC compared with non-HCC patients, but did not reach statistical significance. The preS2:F22L mutation was found in genotypes A and D.

CONCLUSION: Kerala is the first Indian state in which subgenotype A1 has been found to predominate in liver disease patients who developed HCC at a relatively young age.

© 2013 Baishideng Publishing Group Co., Limited. All rights reserved.

Key words: Hepatocellular carcinoma; Cirrhosis; Chronic hepatitis; Phylogenetic analysis; Genotype; India

Core tip: This study shows the predominance of subgenotype A1 in liver disease patients in Kerala, and its high prevalence in hepatocellular carcinoma (HCC) patients. Subgenotype A1 could be more hepatocarcinogenic and HCC could develop at an earlier age, regardless of host ethnicity. The open reading frame of subgenotype A1 isolates from Kerala clustered separately within the “Asian” cluster and encoded distinct subgenotype A1 amino acids. A higher frequency of G1862T was detected compared with subgenotype A1 isolates from other geographical regions. This is the first time that preS deletion mutants have been described in Indian HCC patients.

INTRODUCTION

Hepatitis B virus (HBV) is the prototype member of the family *Hepadnaviridae*. HBV replicates by reverse transcription using a polymerase that lacks proof reading ability, and sequence heterogeneity is a feature of this virus. Phylogenetic analysis of HBV full-length genomes has led to the classification of HBV into nine genotypes (A-I), defined by an intergroup divergence in the complete HBV genome sequence of 7.5\% or more. A tenth genotype J, which was found in a single individual, has been proposed\[9\]. Genotypes A, B, C, D, F and I are further classified into subgenotypes. Most genotypes, and some subgenotypes, display distinct geographical distributions. Moreover, HBV genotypes and, in some cases, subgenotypes, have been shown to play an important role in the clinical consequences of the infection, as well as in the response to antiviral treatment.

HBV infection remains a significant global health problem, with an estimated two billion people infected and more than 240 million chronic carriers of the virus, leading to 600000 deaths from the clinical consequences of infection, including cirrhosis, liver failure and hepatocellular carcinoma (HCC). With a population of more than 1.2 billion people, India has the second largest global pool of chronic HBV infection and HBV is the major cause of liver disease in India\[5\].

Most studies have estimated the hepatitis B surface antigen (HBsAg) carrier rate to be between 2\% and 8\%, placing India within the zone of intermediate endemicity. An HBsAg prevalence rate of 2.97\% was found among the rural population\[3\], and a meta-analysis has reported the mean prevalence in the general population of India as 3.3\%\[4\]. However, these estimates have been questioned because, according to Phadke and Kale\[5\], the often quoted estimate for India of 4.7\% was obtained by incorrectly pooling results of a set of studies including unrepresentative high risk groups and also equating the single test HBsAg positivity rate with the carrier rate. By correcting for these errors, they estimated a carrier rate of 1.4\%.

The known HBV genotype distribution in India is summarized in Figure 1. Overall, at approximately 65\%, genotype D predominates, being the dominant genotype in Delhi in the north, Pune in the west and the Nicobar Islands in the south. Genotype A has been found in approximately 30\%, with the highest frequency found in northern India. At approximately 5\%, genotype C is found in the minority, with the highest frequency in eastern and southern India. The subgenotypes that have been described in India include A1, A2, C1, C2, D1, D2, D3, D5 and D9\[6-9\].

Kerala is the most densely populated state of India, with a population of 33 million. HBsAg prevalence of 0.5\% in the normal population has been reported in northern Kerala\[10\] and an HBsAg prevalence of 1.5\% was detected among voluntary blood donors from Trivandrum, South Kerala\[11\]. There is a paucity of information on the prevalence of HBV genotypes and the respective subgenotypes in Kerala, as well as their association, if any, with different clinical manifestations following infection with HBV. We investigated the distribution of HBV genotypes/subgenotypes among patients with different clinical manifestations of HBV infection and characterized the viral isolates molecularly.

MATERIALS AND METHODS

Patients

The cross-sectional study was conducted from January 2005 to December 2009 during which sera and clinical data were collected from 91 patients diagnosed with chronic HBV infection and HBV-related HCC from
Medical College Trivandrum, Kerala, India. The serum samples were stored at -80°C until use. A serum alanine transaminase (ALT) level of < 10 times the upper limit of normal (ULN), a serum bilirubin level of less than 2.5 times the ULN and detectable HBsAg for ≥ 6 mo were used as inclusion criteria. The presence of the hepatitis B e antigen (HBeAg) was examined at the time of screening. All patients were negative for antibodies to hepatitis C virus, hepatitis D virus and human immunodeficiency virus. The study protocol conformed to the 1975 Declaration of Helsinki. The ethics committees of the Medical College Trivandrum, India and the University of the Witwatersrand, South Africa approved the study.

The diagnosis of HBV-related liver disease was based on clinical data, laboratory tests, liver biopsy and imaging studies. The patients were classified into three groups: group-I (HCC): the 44 patients with HCC were diagnosed by ultrasound scan and elevated serum α-fetoprotein levels (≥ 400 ng/mL) and the presence of a lesion of ≥ 5 cm; group-II (CR-Cirrhosis): 22 patients, with necro-inflammatory damage, fibrosis with nodule formation confirmed by liver biopsy, and with ultrasonographical evidence of portal hypertension; group-III (CH-Chronic Hepatitis): 25 patients, with HBsAg positive status for ≥ 6 mo with normal or intermittently elevated ALT (1.5 times the ULN). Patients in this group were considered for liver biopsy on the basis of elevated ALT levels and HBeAg-status, and diagnosed with cirrhosis using histological activity index (HAI) and Fibrosis scores.

Serological assays
All serum samples were screened for HBsAg and HBeAg using enzyme linked immunosorbent assay kits (DiaSorin S.P.A, Italy), according to the manufacturer’s instructions. Laboratory evaluation included routine liver biochemistry (ALT and aspartate transaminase levels), total bilirubin, albumin, alkaline phosphatase, total protein and prothrombin time. Liver function tests were performed to find necro-inflammatory activity using a Hitachi 902 Fully Automated Chemistry Analyzer (Hitachi High-Technologies Corporation, Tokyo, Japan). The ULN of ALT (40 IU/L) was used for diagnosis.

Real-time polymerase chain reaction quantification of HBV DNA
Polymerase chain reaction (PCR) primers, HBV-Taq1
fragments

The complete S ORF was analyzed as three overlapping sequences. The BigDye Terminator v3.0 Cycle Sequencing Ready Reaction Kit and sequencing was performed with the ABI 3130xl Genetic analyzer (Applied Biosystems). Primers S1F 5'-CAATCGCCGCGTCCGAGGA- CATCTCACTC-3' (2410-2439 from the EcoRI site) and SIR 5'-TCCAGAGGCGCTGGAGGAAAACA-3' (1314-1291 from the EcoRI site) were used for the first round PCR using previously reported reaction conditions. The samples that did not amplify in the first round PCR were used for the second round PCR using previously reported reaction conditions. The samples that did not amplify in the full S region were genotyped using restriction fragment length polymorphism (RFLP) assay for genotyping and molecular characterization.

PCR and restriction fragment length polymorphism assay for genotyping and molecular characterization

Total HBV DNA was extracted from serum using a QIAamp DNA Blood Mini kit (QIAGEN GmbH, Hilden, Germany), according to the manufacturer's instructions. The complete S open reading frame (ORF) was amplified using nested PCR.

Primers S1F 5'-CAATCGCCGCGTCCGAGGA- CATCTCACTC-3' (2410-2439 from the EcoRI site) and SIR 5'-TCCAGAGGCGCTGGAGGAAAACA-3' (1314-1291 from the EcoRI site) were used for the first round PCR using previously reported reaction conditions. The samples that did not amplify in the full S region were genotyped using restriction fragment length polymorphism (RFLP) assay for genotyping and molecular characterization. Subgenotypes of A were also determined using a previously described RFLP assay, which uses the StuI recognition site, 5' AGG↓CCT3' at position 967-972 from the EcoRI site, found only in subgenotype A2 and genotype D, but not in subgenotype A1 [18,19]. Thus, subgenotypes A1 and A2 could be differentiated. The basal core promoter (BCP)/Pre C region of HBV isolates was amplified using nested PCR.

Primers BCP1F 5'-GCATGGAGACCACCGT- GAAC-3' (1606-1625 from the EcoRI site) and BCP1R 5'-GGAAAGAGTCCGAGGCA-3' (1974-1955 from the EcoRI site) were used for the first round and BCP2F 5'-ATGTTAGTATCCCTTGACT- CATAGTGGG-3' (2451-2482 from the EcoRI site) and BCP2R 5'-AGTCCGAGCCATGATGGCAGAGGA-3' (1280-1254 from the EcoRI site) were used for the second round PCR using previously reported reaction conditions. The lower detection limit of our assay was approximately 20 IU/mL. The conversion formula of IU = copies/4.7 was used [12].

Phylogenetic analysis

The complete surface DNA sequences were assembled and aligned manually using MEGA 5 (http://www.megasoftware.net/mega.php). The sequences were compared with the sequences of HBV of genotypes (A-J) from GenBank. The evolutionary history was inferred using the neighbor-joining method and the evolutionary distances computed using the Kimura 2-parameter method. Bootstraping was performed using 1000 replicates to determine the support for the specific nodes. The accession numbers of HBV isolates sequenced in this study have been deposited in GenBank as KC752137-KC752206.

Statistical analysis

Data were represented as mean ± SD. Continuous variables were compared using an independent Student's t test. The χ² test or Fisher's exact test was used to compare categorical variables. Odds ratio was calculated to assess the risk of HCC. All P values were two sided, and the difference was considered statistically significant for P < 0.05. The analysis was performed using Statistical package for Social Sciences (SPSS 15) program (SPSS Inc., Chicago, IL, United States).

RESULTS

Genotyping and phylogenetic analyses of HBV isolated from liver disease patients

Of the 91 HBsAg-positive sera, 86 were successfully genotyped using either RFLP or phylogenetic analysis of the S region (Table 1). Using the Lindh RFLP assay [16] for 36 HBV isolates, 30 belonged to genotype A and six to genotype D. Of the 30 genotype A isolates, 28 were subgenotype A1 and two were subgenotype A2, as determined by an alternative RFLP [16].

Following phylogenetic analysis of the complete S ORF of 50 isolates, 32 belonged to genotype A (subgenotype A1:A2, 31:1) (Figure 2A), 17 to genotype D (subgenotypes D1:D2:D3, 4:12:1) (Figure 2B) and one to genotype C (subgenotype C1) (Figure 2A). The genotype A strains belonged to serotype adw2 (84.4%) and ayw1 (15.6%). The genotype D strains were of serotype ayw3 (58.8%), ayw2 (35.3%) and adw3 (5.9%). The single subgenotype C1 strain was adw.

The subgenotype A1 isolates split into an “African” and an “Asian” cluster [10] (Figure 2A). The 31 subgenotype A1 isolates from Kerala clustered within the Asian clade as a separate monophyletic clade and encoded the distinct subgenotype A1 amino acids, preS1:Q54, preS1:V74, preS1:A86, and preS1:V91 in the preS1 region and preS2: L32 in the preS2 region [17]. The majority of the isolates in the “Asian” cluster, including the Kerala isolates, had preS1:S5, preS1:S6, preS1:F25. The isolates in the African clade displayed greater variation, with preS1:S5, preS1:6; preS1:S5, preS1:A6 or preS1:51, preS1:6P. There were, however, a number of amino acids in the preS1 and preS2 regions that differentiated the Kerala clade
from other Asian strains. The majority of Kerala strains (22/31; 71%), had preS1:N48 in the preS1, whereas V1/ N/T was found in the other clades. In the preS2, 28/31 (90%), had preS2:T7, whereas the other Asian strains had either preS2:T7 or preS2:A7. In contrast to the other Asian strains that had preS2:T37, the Kerala strains had preS2:N37, as did the strains in the African clade. PreS2: P54 in the preS2 was found in 90% of the Kerala strains, whereas the other clades had a higher diversity of amino acids at this position (Figure 2). A cut off of 60% amino acid sequence identity was used to define the consensus sequence within the clades.

The Keralite genotype D isolates had the characteristic 33-nucleotide deletion in the preS1 region and had a relatively well-conserved polymerase overlapping preS1/preS2/S region when compared to subgenotype A1 sequences. The preS2 signature amino acids, preS2: I42, preS2:S43 and preS2:L54, were found in the Keralite strains belonging to subgenotype D1; however, they differed by having preS2:39V instead of preS2:39A.

### Demographic, clinical and virological characteristics

In all three disease groups, the frequency of males was significantly higher (Table 1). Patients with HCC were significantly older than CH patients ($P = 0.0001$). Twenty seven percent of the whole cohort was HBeAg-positive, with no significant difference in the frequency between the three disease groups. HBeAg-positive individuals were significantly younger than HBeAg-negative ($32.1 ± 17.9$ years vs $39.6 ± 11.1$ years, $P = 0.032$) and had higher viral loads ($5.4 ± 1.8$ log$_{10}$ IU/mL vs $3.6 ± 1.6$ log$_{10}$ IU/mL, $P = 0.016$). The ALT levels differed significantly between HBeAg-positive and negative patients ($32.3$ IU/L vs $37.4$ IU/L, $P = 0.012$, equal variances not assumed). HCC patients had higher viral load defined by HBV DNA level $\geq 4.7 \times 10^4$ IU/mL compared with the non-HCC patients ($18/29(62.1\%)$ vs $10/33 (30.3\%)$, respectively, $P = 0.012$, OR $= 3.76$, 95%CI: 1.16–12.52). The mean ALT values did not vary significantly between the three disease groups.

The majority of patients (72%) were infected with HBV genotype A, 27% with genotype D and one patient was infected with genotype C. The majority of genotype A strains (95%) belonged to subgenotype A1, and three to A2. Compared with the other disease groups, HCC patients were predominantly infected with subgenotype A1 ($P < 0.05$). Subgenotypes D1, D2 and D3 were found, with D2 (70.6%) predominating followed by D1 (23.5%) and a single strain of D3. Age, HBV viral load, the frequency of HBeAg-positivity and ALT levels, did not differ between those patients infected with genotype A and D in all the three groups. However, HBeAg-negative HCC patients, infected with genotype A, were significantly younger (44.1 ± 8.0 years) than those infected with genotype D (53.0 ± 8.8 years) ($P = 0.02$). There was no significant difference in the mean HAI ($5.8 ± 2.8$ vs $4.6 ± 3.1$) and fibrosis scores ($1.0 ± 1.8$ vs $2.0 ± 2.0$) between those with genotypes A and D. Genotype A was seen in 70% of the patients with HAI $\geq 4$.

Detection of mutations in the BCP/Pre C region was performed for 63 HBV isolates (Table 2). BCP and/or Pre C mutants were detected in 63% of the isolates, with different mutational patterns found in genotypes A and D (Figure 3). Mutation 1773T was characteristic of genotype A and 1773C of genotype D, with no significant difference in the frequency of the mutation in the disease groups. Mutation G1896A occurred only in genotype D, whereas C1766T and G1862T occurred more frequently in genotype A (Figure 3). The double mutation A1762T/ G1764A was found in 26 (41.2%) isolates, with the single mutation, G1764A, occurring in 12 (19%) isolates. Both the single and double BCP mutations were significantly

### Table 1 Demographic, clinical and virological characteristics of hepatitis B surface antigen-positive patients with different disease profiles

| Characteristic | Group I HCC ($\eta = 44$) | Group II CR ($\eta = 22$) | Group II CH ($\eta = 25$) | Total ($\eta = 91$) |
|---------------|----------------|----------------|----------------|----------------|
| Demographic and clinical data | Gender (M/F) | 33/11 | 18/4 | 18/7 | 69/22 |
| Age (yr) (mean ± SD) | 48.70 ± 10.94 | 40.68 ± 11.52 | 26.28 ± 11.10 | 40.87 ± 14.66 |
| ALT, IU/L | 73.50 ± 11.65 | 70.50 ± 41.09 | 56.32 ± 30.01 | 67.45 ± 37.74 |
| HBeAg positive | 7 (24.4) | 4 (14.3) | 9 (37.5) | 20 (26.67) |
| Virological characteristics | HBV DNA log$_{10}$ copies/mL | 4.79 ± 1.41 | 3.38 ± 1.69 | 3.27 ± 2.12 | 4.03 ± 1.83 |
| Number genotyped by direct sequencing/RFLP | Genotype A | 33 (82.5) | 14 (66.6) | 15 (60) | 62 (72.1) |
| | Subgenotype A1 | 30 | 14 | 15 | 59 (95) |
| | Subgenotype A2 | 3 | - | - | 3 (5) |
| | Genotype C | - | 1 (4.8) | - | 1 (1.2) |
| | Genotype D | 7 (17.5) | 6 (28.6) | 10 (40) | 23 (26.7) |
| | Subgenotype D1 | - | 1 | 3 | |
| | Subgenotype D2 | 4 | 2 | 6 | |
| | Subgenotype D3 | 1 | - | - | |

*P < 0.05 vs chronic hepatitis. 1Depletion of serum allowed the viral loads to be determined for only 28, and hepatitis B e antigen for 29 HCC sera; 2Subgenotyping of genotype D was performed by direct sequencing. CH: Chronic hepatitis; HCC: Hepatocellular carcinoma; CR: Cirrhosis; SD: Standard deviation; ALT: Alanine aminotransferase; RFLP: Restriction fragment length polymorphism; HBeAg: Hepatitis B e antigen.

*"The majority of patients (72%) were infected with HBV genotype A, 27% with genotype D and one patient was infected with genotype C. The majority of genotype A strains (95%) belonged to subgenotype A1, and three to A2. Compared with the other disease groups, HCC patients were predominantly infected with subgenotype A1 (P < 0.05). Subgenotypes D1, D2 and D3 were found, with D2 (70.6%) predominating followed by D1 (23.5%) and a single strain of D3. Age, HBV viral load, the frequency of HBeAg-positivity and ALT levels, did not differ between those patients infected with genotype A and D in all the three groups. However, HBeAg-negative HCC patients, infected with genotype A, were significantly younger (44.1 ± 8.0 years) than those infected with genotype D (53.0 ± 8.8 years) (P = 0.02). There was no significant difference in the mean HAI (5.8 ± 2.8 vs 4.6 ± 3.1) and fibrosis scores (1.0 ± 1.8 vs 2.0 ± 2.0) between those with genotypes A and D. Genotype A was seen in 70% of the patients with HAI ≥ 4.

Detection of mutations in the BCP/Pre C region was performed for 63 HBV isolates (Table 2). BCP and/or Pre C mutants were detected in 63% of the isolates, with different mutational patterns found in genotypes A and D (Figure 3). Mutation 1773T was characteristic of genotype A and 1773C of genotype D, with no significant difference in the frequency of the mutation in the disease groups. Mutation G1896A occurred only in genotype D, whereas C1766T and G1862T occurred more frequently in genotype A (Figure 3). The double mutation A1762T/G1764A was found in 26 (41.2%) isolates, with the single mutation, G1764A, occurring in 12 (19%) isolates. Both the single and double BCP mutations were significantly.
African
preS1: (S5, S6)/(L5, P6), F25, V/I/N/T48
preS2: A7, N37, P/L/W/A/S54
sp: i9, L10

Asian
preS1: (S5, S6), F25, V/I/N48
preS2: A/T7, T37, P/T/S54
sp: i9, L10

Kerala
preS1: (S5, S6), F25, V548
preS2: A/T7, T37, P/T/S54
sp: i9, L10

Genotype A
A1
A2

Gopalakrishnan D et al. Hepatitis B virus infection in Kerala
Gopalakrishnan D et al. Hepatitis B virus infection in Kerala
associated with HCC in both genotypes A and D ($P = 0.03$). Mutation C1766T/T1768A was significantly associated with HCC and found predominantly in subgenotype A1 ($P = 0.05$) (Table 2). Although G1862T was significantly associated with subgenotype A1, occurring in 85%, there was no significant difference between its presence in HCC and non-HCC patients (Table 2) and between isolates from HBeAg-positive and -negative patients (38.5% vs 61.5%, respectively; $P = 0.08$). However, in combination with A1762T/G1764A, G1862T was significantly associated with HCC in patients infected with subgenotype A1 ($P = 0.0004$). There was no correlation between the presence of BCP/Pre C mutations with either the age, gender, or the viral loads.

PreS deletion mutants, whose patterns are depicted in Table 3, were detected in nine isolates from five HCC and four CH, but in none of the CR patients. Overall, seven different types of preS mutations were detected (Table 3). The mean age of the patients, with and without preS deletions, did not differ significantly (35.4 ± 11.5 years vs 37.0 ± 15.8 years, $P = 0.78$), nor did the HBV DNA (3.8 ± 2.0 vs 4.2 ± 1.7, $P = 0.62$) and mean ALT levels (58.2 ± 20.8 vs 60.0 ± 28.2, $P = 0.88$). A higher frequency of preS deletion mutants was observed in HCC patients infected with genotype A, although this did not reach statistical significance (Table 2). The preS2 start codon M1T/I mutation, was unique to genotype A strains, occurring in 5/32 (15.6%) isolates from all disease groups

**Table 2 Multiple logistic regression analysis of basal core promoter/precore region in different disease groups**

|                      | Gen A ($n = 20$) | Gen D ($n = 6$) | Gen A ($n = 26$) | Gen D ($n = 11$) | OR (95%CI) | $P$ value |
|----------------------|-----------------|----------------|-----------------|-----------------|------------|-----------|
| Basic core promoter/ pre core region |                 |                |                 |                 |            |           |
| A1762T/G1764A + G1764A only | 16 (80)         | 5 (83)         | 14 (54)         | 3 (27)          | 20.2 (6.3-65)\(a\) | 0.0082 NS |
| C1766T/T1768A         | 9 (45)          | 1 (5)          | 5 (25)          | 0               | 25 (7.3-86)\(b\) | 0.0301 |
| 1773T (genotype A)    | 19 (95)         | -              | 25              | 7               | 14.3 (1.7-119)\(c\) | 0.0525 |
| 1773C (genotype D)    | -               | 4              | 25              | -               | NS         |
| G1862T                | 18 (90)         | 2 (33)         | 21 (81)         | 1 (9)           | 30.33 (5.62-192.6)\(d\) | 0.00014 |
| G1896A                | 0               | 4 (67)         | 0               | 2 (18)          | -          | NS        |
| A1762T/G1764A + G1862T| 13 (65)         | 1 (2)          | 13 (50)         | 0               | 4.81 (0.6-39.4)\(e\) | 0.0002 |
| Complete S region     |                 | 5 (33.3)\(f\) | 0               | 3 (17.6)        | 1.89 (0.54-6.60) | 0.00045 |

\(a\)Comparison between hepatocellular carcinoma (HCC) and non-HCC, all genotypes; \(b\)Comparison between HCC and non-HCC, restricted to genotype A isolates; \(c\)Comparison between HCC and non-HCC, restricted to genotype D isolates; \(d\)Comparison between genotype A and D isolates; \(e\)Percentage out of 15 genotype A HCC isolates that were sequenced in the S region. CR: Cirrhosis; CH: Chronic hepatitis.

Figure 2 Phylogenetic relationships among complete preS1/pre S2/S sequences (nt 2854-835 numbering according to GenBank accession AY233274). A: Subgenotype A1 from hepatitis B virus (HBV) positive patients from Kerala (marked in bold) compared with sequences obtained from GenBank established using the neighbor joining method; B: Genotype D isolates from HBV positive patients from Kerala (marked in bold) compared with HBV isolates obtained from GenBank established using the neighbor joining method. Bootstrap statistical analysis was performed using 1000 replicates. Each sequence obtained from GenBank is designated by its accession number and country of origin. The characteristic amino acids in the preS1 and polymerase spacer regions are indicated next to the sequences or relevant clades.
Gopalakrishnan D et al. Hepatitis B virus infection in Kerala

and occurred at a higher frequency in isolates from HCC patients ($P = 0.076$). The preS2: F22L/I mutation was detected in 13 isolates (genotype A-9/32, 28%; genotype D-4/17, 23%). The F22 mutation was significantly associated with HCC (10/13, 77%) compared with CR (2/13, 15%) and CH (1/13, 8%), respectively ($P = 0.0065$). This significance remained when comparing genotype D isolates only, but not genotype A isolates alone.

**DISCUSSION**

The present study demonstrated that genotype A was the most prevalent HBV genotype infecting liver disease patients in Kerala. This prevalence differed from other geographical regions of India, where genotype D predominates or occurs at an equal prevalence with genotype A (Figure 1). Ninety five percent of the genotype A isolates belonged to subgenotype A1, which has also been found to be the predominant subgenotype of A in India$^{[10]}$.

The Keralite subgenotype A1 strains clustered with the Asian subgenotype strains but differed from them in some molecular characteristics in the preS2 region, which they shared with African strains. A minority of genotype A strains belonged to subgenotype A2, which was previously described to be restricted to the peripheral blood lymphocytes of eastern Indian blood donors$^{[19]}$. Subgenotype A1 of HBV was the first subgenotype to be recognized and is the dominant genotype A strain in Africa, with unique molecular characteristics that differentiate it from A2, the genotype A strain prevailing outside Africa. Subgenotype A1 has its origin in Africa and its

---

**Table 3  Summary of the pre-S mutations prevalent among the three clinical groups and genotypes**

| Isolate | Age/sex | Clinical status | Subgenotype | Start codon | PreS1 | Nucleotide from the EcoR1 start | PreS2 | Amino acids from preS2 start codon | Functions affected |
|---------|---------|-----------------|--------------|-------------|-------|---------------------------------|-------|-----------------------------------|-------------------|
| CHB42   | 40/M    | CH              | A1           | ATG         | AGC   | -                               | -     | -                                 | A                 |
| HCC12   | 60/M    | HCC             | A1           | ATG         | AGC   | -                               | -     | -                                 | A                 |
| HCC25   | 50/M    | HCC             | A1           | ATG         | AGC   | -                               | -     | -                                 | A                 |
| CHB137  | 16/F    | CH              | A1           | ATG         | ATG   | 18                             | 2854-2871 | 6 | 1-6                               | I                 |
| CHB170  | 29/F    | CH              | D2           | ATG         | ATG   | 21                             | 35-55  | 7 | 16-22                             | T, B, P           |
| CHB202  | 33/M    | CH              | A1           | ATG         | ATG   | 24                             | 28-51  | 8 | 13-21                             | T, B, P           |
| CHB413  | 35/M    | CH              | A1           | ATG         | ATG   | 6                              | 49-54  | 2 | 21-22                             | B, P              |
| HCC29   | 50/F    | HCC             | A1           | ATG         | ATA   | 33                             | 22-54  | 11 | 11-22                             | A, T, B, P        |
| HCC30   | 32/M    | HCC             | A1           | ATG         | ATG   | 24                             | 30-53  | 8 | 13-21                             | T, B, P           |
| HCC37   | 30/F    | HCC             | A1           | ATG         | ATG   | 33                             | 24-56  | 11 | 11-22                             | T, B, P           |
| HCC59   | 55/M    | HCC             | A1           | ATG         | ATG   | 33                             | 24-56  | 11 | 11-22                             | T, B, P           |
| HCT3    | 38/M    | HCC             | A1           | ATG         | ATA   | 54                             | 1-54   | 18 | 4-22                              | A, T, B, P, M     |

A: PreS2 initiation codon abolished; MIT, MII; I: PreS1 initiation start codon abolished; M: Morphogenesis domain ps1:103-119 and ps2: 1-13; B: B cell epitope, amino acids 14-26; P: Putative neutralizing anti-preS2 antibody, amino acids 1-26; HCC: Hepatocellular carcinoma; CHB: Chronic hepatitis B.
global dispersal coincides with historical events, including the slave trade and colonization. Calicut and Cochin on the west coast of Kerala were major sea ports frequented by both the Dutch East India Company and Portuguese colonists.

The patients infected with genotype A and D did not differ from each other in terms of age, HBV viral loads, the frequency of HBeAg-positivity and ALT levels (Table 1). However, compared to cirrhotic and chronic hepatitis patients, a significantly higher proportion of HCC patients were infected with subgenotype A1, and HBsAg-negative HCC patients infected with subgenotype A1 were significantly younger than HCC patients infected with genotype D. Although the number of HCC patients in the present study is relatively low, the results concur with a South African study that showed that patients infected with subgenotype A1 had a 4.5 fold increased risk of developing HCC compared with those infected with non-A genotypes and they developed the cancer 6.5 years earlier, thus intimating a higher hepatocarcinogenic potential of subgenotype A1, regardless of host ethnicity. However, studies with a larger number of HCC patients would be required to confirm this. These findings differ from a New Delhi study, which found a comparable distribution of genotype A and D between disease groups and that genotype D, and not genotype A, was associated with HCC and were of younger age. Other studies from New Delhi and Western India found no association between genotype A, D and disease progression. The subgenotypes of A were not differentiated in any of these three studies.

Irrespective of the genotype, the frequency of the A1762T/G1764A and 1764A mutations was significantly higher in HCC patients compared with non-HCC patients. The majority of HCC patients were infected with subgenotype A1. Similarly, the BCP double mutation occurred at a higher frequency in HCC patients compared with asymptomatic carriers in southern Africa, where subgenotype A1 predominates. This was not the case in Western India where genotype D is prevalent. Although the BCP mutants have been reported to contribute to the HBeAg-negative phenotype by downregulating precore mRNA transcription, in the present study, and in agreement with others, there was no correlation between the presence of the A1762T/G1764A mutation and HBsAg-positivity.

The double mutation C1766T/T1768A was significantly associated with HCC and subgenotype A1 (Table 2). The T1768A mutation results in F132Y in HBx and may play a synergic role with K130M and V131I, introduced by A1762T/G1764A, leading to carcinogenesis. Moreover, mutation C1766T/T1768A has been reported as an independent predictor of cirrhosis in HBsAg-negative patients and is associated with higher viral replication by increasing the encapsidation of pg RNA.

Mutation G1862T was found in 85% of subgenotype A1 isolates (Table 2). Previously, G1862T was detected in 79% of global subgenotype A1 isolates, but in none of the subgenotype A2 isolates, and has been shown to be a characteristic of subgenotype A1. This high frequency of G1862T in the Keralite strains is much higher than that reported in either Eastern Indian and Southern African studies (60% and 250%, respectively). In the present study, the combination of A1762T/G1764A and G1862T was significantly associated with HCC in patients infected with subgenotype A1. A previous study showed that G1862T was significantly associated with HBsAg-negativity in South African HCC patients, but not in asymptomatic carriers. Moreover, mutation A1762T/G1764A is found frequently in South African HCC patients, but not in asymptomatic carriers, and the majority of South African HCC patients are infected with subgenotype A1.

The present study showed a significant association of the preS2:F22L mutation with the development of HCC, particularly in genotype D. Recent studies identified the F22L mutation in the preS2 region as a risk factor for HCC among patients infected with genotype C, and showed significant association of this mutation with liver cirrhosis in Eastern India. Our study supports the possibility that F22L may be associated with severe liver disease progression.

The preS deletion and initiation codon mutations were prevalent in strains isolated from all clinical groups (Table 3). This is the first study to describe the preS mutants from Indian HCC patients infected with subgenotype A1. A strong correlation between preS mutants and the development of HCC has been shown in patients infected with genotypes B or C. In studies carried out in isolates belonging to genotypes B and C, mutated envelope proteins were shown to accumulate within the hepatocyte endoplasmic reticulum (ER) and result in a characteristic histopathological hallmark of HCC, known as ground glass hepatocytes. HBV induced ER stress has been shown to dysregulate several cell cycle regulatory pathways, which may contribute to hepatocarcinogenesis.

In conclusion, genotypes A and D were isolated from liver disease patients in Kerala, Southern India, with subgenotype A1 predominating. The relatively high prevalence of subgenotype A1 in HCC patients supports previous studies in Africa, which showed an association of subgenotype A1 with HCC and its development at a younger age. This association appears not to depend on host ethnicity. The combination of BCP and/or preC mutations, as well as the described preS mutations, could lead to the accumulation of replicative intermediates and viral proteins, contributing to viral integration and cellular stress or damage. These combined characteristics could induce severe liver disease, including HCC, and should be explored further.

ACKNOWLEDGMENTS

We thank Professor Michael Kew, University of Cape Town, South Africa for initiating the collaboration between Professors Shenoy KT of India and Kramvis A of South Africa.
Background

Genotype D, followed by genotype A, has been reported to predominate in India. Hepatitis B virus (HBV) isolates with the A1762T/G1764A variations are prevalent in hepatocellular carcinoma (HCC) patients. In addition, variation G1862T has been reported to be a characteristic of subtype A1. Subgenotype A1 has been reported to be more hepatocarcinogenic in southern Africans, who develop HCC at a younger age than those infected with other genotypes.

Research fronts

The hepatocarcinogenic potential of subgenotype A1 of HBV has been linked to disease progression in regions outside India. Subgenotype A1 has its origin in Africa and its global dispersal coincides with historical events, including the slave trade and colonization.

Innovations and breakthroughs

This is the first study to report the predominance of subgenotype A1 in liver disease patients in India and its high prevalence in HCC patients. The relatively high prevalence of subgenotype A1 in HCC patients supports previous studies in Africa that showed an association of subgenotype A1 with HCC and its development at a younger age. The open reading frame of subgenotype A1 isolates from Kerala clustered within the Asian clade as a separate clade and encoded distinct subgenotype A1 amino acids. The subgenotype A1 isolates from Kerala had a higher frequency of G1862T compared to subgenotype A1 isolates from other geographical regions. This is the first time that preS deletion mutants have been described in Indian HCC patients. Pre-S2: F2ZL was found in genotypes A and D.

Applications

The prevalence of different mutations in the various genotypes of HBV may serve as biomarkers for disease risk and development of HCC. The differences in the geographical distribution of genotypes and subgenotypes of HBV may require different treatment algorithms. Knowledge of HBV genotypes and/or mutations may facilitate personalized treatment.

Terminology

A genotype is generally defined as the genetic constitution of an organism. In the case of viruses, the term genotype applies to the forms into which the genome sequence has stabilized after a prolonged period of time and that are replication competent. The genotypes of HBV are defined by an intergroup divergence of more than 7.5%-8% in the complete genome sequence and by genomic sequence has stabilized after a prolonged period of time and that are

Peer review

The study contributes to the understanding of the relationship between HBV variability and clinical outcomes in different populations worldwide. The relationship between the preS deletion and HCC in this study is a critical and hot point.

References

1. Tamatsuki K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M. 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: 10538-10547 [PMID: 19640977 DOI: 10.1128/ JVI.00462-09]

2. Acharya SK, Madan K, Dattagupta S, Panda SK. Viral hepatitis in India. Natl Med J India 2006; 19: 203-217 [PMID: 17100109]

3. Chowdhury A, Santra A, Chakravarty R, Banerji A, Pal S, Dhali GK, Datta S, Banerji S, Mani B, Chowdhury SR, Bhattacharya SK, Mazumder DG. Community-based epidemiology of hepatitis B virus infection in West Bengal, India: prevalence of hepatitis B e antigen-negative infection and associated viral variants. J Gastroenterol Hepatol 2005; 20: 1712-1720 [PMID: 16246191 DOI: 10.1111/ j.1440-1746.2005.04070.x]

4. Thiyagarajan SP, Murugavel KG. Prevalence of hepatitis B virus infection in general population of India. Ind J Gastroenterol 2000; 19 (Suppl 3): C110

5. Phadke A, Kale A. HBV carrier rate in India. Indian Pediatr 2002; 39: 787-788; author reply 788-790 [PMID: 12196697]

6. Vivekanandan P, Abraham P, Sridharan G, Chandy G, Daniel D, Raghuraman S, Daniel HD, Subramaniam T. Distribution of hepatitis B virus genotypes in blood donors and chronically infected patients in a tertiary care hospital in southern India. Clin Infect Dis 2004; 38: e81-e86 [PMID: 15127358 DOI: 10.1086/383144]

7. Youssif M, Kramvis A. Genotype D of hepatitis B virus and its subgenotypes: An update. Hepatol Res 2013; 43: 355-364 [PMID: 22987460 DOI: 10.1111/j.1872-034X.2012.01090.x]

8. Ghosh S, Banerjee P, Deny P, Mondal RK, Nandi M, Roychoudhury A, Das K, Banerjee S, Santra A, Zoulim F, Chowdhury A, Datta S. New HBV subgenotype D9, a novel D/C recombinant, identified in patients with chronic HBsAg-negative infection in Eastern India. J Viral Hepat 2013; 20: 209-218 [PMID: 23383660 DOI: 10.1111/j.1365-2893.2012.01655.x]

9. Banerjee A, Datta S, Chandra PK, Roychoudhury S, Panda CK, Chakravarty R. Distribution of hepatitis B virus genotypes: phylogenetic analysis and virological characteristics of genotype C circulating among HBV carriers in Kolkata, Eastern India. World J Gastroenterol 2006; 12: 5964-5971 [PMID: 17009394]

10. Sandesh K, Varghese T, Harikumar R, Beena P, Sasidharan VP, Bindu CS, Tony J, Harish K, Sunilkumar K, Ramachandran TM. Prevalence of Hepatitis B and C in the normal population and high risk groups in north Kerala. Trop Gastroenterol 2006; 27: 80-83 [PMID: 17089617]

11. Anjali H, Issac A, Anjali MR, Anish TS. Transfusion-transmissible infections among voluntary blood donors at Government Medical College Thrissurvananthapuram, Kerala, India. Asian J Transf Sci 2012; 6: 55-56 [PMID: 22623852 DOI: 10.4103/0973-6472.95960]

12. Finnhaber C, Viana R, Reynke A, Schultz D, Malope B, Maskew M, Die Biscieglie A, MacPhail P, Sanne I, Kew M. Occult hepatitis B virus infection in patients with isolated core antibody and HIV co-infection in an urban clinic in Johannesburg, South Africa. Int J Infect Dis 2009; 13: 488-492 [PMID: 19081280 DOI: 10.1016/j.ijid.2008.08.018]

13. Vermeulen M, Dickens C, Lelie N, Walker E, Coleman B. Virus transmission by C, Keyter M, Reddy R, Crookes K, Kramvis A. Hepatitis B blood transfusion during 4 years of individual-donation nucleic acid testing in South Africa: estimated and observed window period risk. Transfusion 2012; 52: 880-892 [PMID: 21981368 DOI: 10.1111/j.1537-2995.2011.03553.x]

14. Lindh M, Andersson AS, Gusdal A. Genotypes, nt 1858 variants, and geographic origin of hepatitis B virus--large-scale analysis using a new genotyping method. J Infect Dis 1997; 175: 1285-1293 [PMID: 9180165]

15. Kew MC, Kramvis A, Yu MC, Arakawa K, Hodkinson J. Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu-speaking sub-saharan Africans. J Med Virol 2005; 75: 513-521 [PMID: 15714494 DOI: 10.1002/jmv.20311]

16. Makondo E, Bell TG, Kramvis A. Genotyping and molecular characterization of hepatitis B virus from human immunoodeficiency virus-infected individuals in southern Africa. PLoS One 2012; 7: e46345 [PMID: 23029487 DOI: 10.1371/ journal.pone.0046345]

17. 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: 1211-1220 [PMID: 15105537]

18. Banerjee A, Kurbanov F, Datta S, Chandra PK, Tanaka Y, Mizokami M, Chakravarty R. Phylogenetic relatedness and genetic diversity of hepatitis B virus isolates in Eastern India. J Med Virol 2006; 78: 1164-1174 [PMID: 16847957 DOI: 10.1002/jmv.20677]

19. Datta S, Panigrahi R, Biswas A, Chandra PK, Banerjee A,
mutants in serum and liver of Southern African Blacks with hepatocellular carcinoma. J Hepatol 1998; 28: 132-141 [PMID: 9537850 DOI: 10.1016/S0168-8278(98)00212-2]

35 Munn HS, Lee SA, Kim H, Hwang ES, Kook YH, Kim BJ. Novel F141L pre-S2 mutation in hepatitis B virus increases the risk of hepatocellular carcinoma in patients with chronic genotype C infections. J Virol 2011; 85: 123-132 [PMID: 20962085 DOI: 10.1128/JVI.01524-10]

36 Biswas A, Panigrahi R, Banerjee A, Pal M, De BK, Chakrabarti S, Chakravarty R. Differential pattern of pre-S mutations/deletions and its association with hepatitis B virus genotypes in Eastern India. Infect Genet Evol 2012; 12: 384-391 [PMID: 22266243 DOI: 10.1016/j.meegid.2012.01.007]

37 Sugauchi F, Ohno T, Orito E, Sakugawa H, Ichida T, Komatsu M, Kuramitsu T, Ueda R, Miyakawa Y, Mizokami M. Influence of hepatitis B virus genotypes on the development of preS deletions and advanced liver disease. J Med Virol 2003; 70: 537-544 [PMID: 12794715 DOI: 10.1002/jmv.10428]

38 Huy TT, Ushijima H, Win KM, Luengojanakul P, Shrestha PK, Zhong ZH, Smirnov AV, Taltavull TC, Sata T, Abe K. High prevalence of hepatitis B virus pre-S mutant in countries where it is endemic and its relationship with genotype and chronicity. J Clin Microbiol 2003; 41: 5449-5455 [PMID: 14662924]

39 Gerner P, Schäfer HM, Prange R, Pravitt D, Wirth S. Functional analysis of a rare HBV deletion mutant in chronically infected children. Pediatr Res 2003; 53: 891-897 [PMID: 12646724 DOI: 10.1203/01.PDR.0000064906.63993.72]

40 Hisieh YH, Su IJ, Wang HC, Chang WW, Lei HY, Lai MD, Chang WT, Huang W. Pre-S mutant surface antigens in chronic hepatitis B virus infection induce oxidative stress and DNA damage. Carcinogenesis 2004; 25: 2023-2032 [PMID: 15180947 DOI: 10.1095/carcin/bgh207]

41 Kumar A, Kumar SI, Pandey R, Naik S, Aggarwal R. Hepatitis B virus genotype A is more often associated with severe liver disease in northern India than is genotype D. Indian J Gastroenterol 2005; 24: 19-22 [PMID: 15778521]

42 Sharma S, Sharma B, Singla B, Chawla YK, Chakraborti A, Saini N, Duseja A, Das A, Dhiman RK. Clinical significance of genotypes and precore/core basal promoter mutations in HBV related chronic liver disease patients in North India. Dig Dis Sci 2010; 55: 794-802 [PMID: 20043209 DOI: 10.1007/s10620-009-1083-1]

43 Arankalle VA, Gandhe SS, Borkakoty BJ, Walimbe AM, Biswas A, Mahanta J. A novel HBV recombinant (genotype I) similar to Vietnam/Laos in a primitive tribe in eastern India. J Virol Hepat 2010; 17: 501-510 [PMID: 2059669 DOI: JHV1206]

44 Duong TN, Horike N, Michitaka K, Yan C, Mizokami M, Tanaka Y, Jyo Yoko K, Yamamoto K, Miyaoka H, Yamashita Y, Ohno N, Onji M. Comparison of genotypes A, C, and D among the same ethnic population. J Med Virol 2011; 83: 253-260 [PMID: 21181919 DOI: 10.1002/jmv.21979]

45 Panigrahi R, Biswas A, Banerjee A, Singh SP, Panigrahi MK, Roque-Afonso AM, Das HS, Mahapatra PK, Chakraborti S, Chakravarty R. Subgenotype D3, BCP and MHR mutations in hepatic complications among hepatitis B virus infected patients from Orissa, India. Infect Genet Evol 2012; 12: 1622-1629 [PMID: 22820088 DOI: 10.1016/j.meegid.2012.06.015]

46 Mukherjee RM, Balkumar Reddy P, Sasikala M, Nagaraja...
Rao P, Nageshwar Reddy D. Identification of genotype B among hepatitis B virus-infected patients in Hyderabad, India. *Ann Hepatol* 2009; 8: 269-270 [PMID: 19841511]

Arankalle VA, Murhekar KM, Gandhe SS, Murhekar MV, Ramdas AY, Padbidri VS, Sehgal SC. Hepatitis B virus: predominance of genotype D in primitive tribes of the Andaman and Nicobar islands, India (1989-1999). *J Gen Virol* 2003; 84: 1915-1920 [PMID: 12810887 DOI: 10.1099/vir.0.18943-0]

Bell TG, Kramvis A. Mutation Reporter Tool: an online tool to interrogate loci of interest, with its utility demonstrated using hepatitis B virus. *Virol J* 2013; 10: 62 [PMID: 23433201 DOI: 10.1186/1743-422X-10-62]

P- Reviewers: Aghakhani A, Panduro A, Rigato I, Zhao P
S- Editor: Gou SX  L- Editor: Stewart GJ  E- Editor: Wang CH
