

Short Report

Open Access

Hepatitis B virus subgenotypes D1 and D3 are prevalent in Pakistan

Saeeda Baig*¹, Anwar Siddiqui², Runu Chakravarty³ and Tariq Moatter⁴

Address: ¹Department of Biochemistry, Ziauddin Medical College, Ziauddin University, Karachi, Pakistan, ²Department of Biological and Biomedical Sciences, Juma Building, Aga Khan University, Karachi, Pakistan, ³ICMR Virus Unit, ID & BG Hospital Campus, Kolkata, West Bengal, India and ⁴Department of Molecular Pathology Juma Building, Aga Khan University, Karachi, Pakistan

Email: Saeeda Baig* - baigsaeeda@yahoo.com; Anwar Siddiqui - anwar.siddiqui@aku.edu; Runu Chakravarty - runugc@hotmail.com; Tariq Moatter - tariq.moatter@aku.edu

* Corresponding author

Published: 4 January 2009

Received: 16 October 2008

BMC Research Notes 2009, 2:1 doi:10.1186/1756-0500-2-1

Accepted: 4 January 2009

This article is available from: <http://www.biomedcentral.com/1756-0500/2/1>

© 2008 Baig et al; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: As the hepatitis B genotyping is important for assessing its clinical implications and geographical distribution, the sub-genotypes have been found useful for determination of specific genomic markers related to hepatocarcinogenesis. In Pakistan, there is no reported data on molecular evolutionary analysis of HBV. A study was, therefore, much needed to evaluate the spectra of mutations present in the strains prevalent here.

Findings: to confirm specificity of PCR typing, phylogenetic analysis of the pre-S1 region and the divergence was studied through 13 sequences of 362 bp (accession number [EF432765](#) – [EF432777](#)). A total of 315 serum samples, selected from HBsAg positive patients representing the major ethnic groups, residing in Karachi, Sindh were tested for genotyping. Genotype D (219/315) was found to be the most prevalent (70%) amongst our patients. The rest of the genotypes A and a mixture of A and D (AD) were distributed as 20%, and 10% respectively. Phylogenetic tree demonstrated clustering of 11 samples with subgenotype D1 sequences and the remaining two strains on a branch within D3 samples. All samples intermixed with strains from other countries and were found to be closely related to Indian, Iranian and Egyptian HBV strains with 98.7 – 99.0% homology.

Conclusion: This study confirms the predominance of genotype D in southeastern Asia and presence of subgenotypes D1 and D3 in the Pakistani infected patients. More studies are required to investigate the reason for fewer inclusions of D3 compared to the D1 in Pakistani HBV strains.

Findings

The structural and functional differences between HBV genotypes are the mainstay to severity, complications, treatment and possibly vaccination against the virus. Sub-genotypes have been identified in different HBV genotypes on the basis of a > 4% (but < 8%) difference in the complete nucleotide sequences [1]. The phylogenetic method, a reliable tool to study the divergence in the HBV, are now being utilized to elucidate the history and origin of HBV genome scattered geographically around

the world [2]. It has become well known that the genotypes/sub-genotypes vary in the geographical areas and correlate strongly with ethnicity [3]. Pakistan is one of the intermediate HBV endemic countries having distinct multiethnic populations. However, sufficient information on the molecular epidemiology of HBV is not available. There is no data on subgenotypes and their recombination or the phylogenetic relatedness of the virus endemic in the country. Karachi, the largest city of Pakistan and capital of the province of Sindh, is considered 'mini Paki-

stan' as all ethnic groups including Punjabis, Sindhis, Pukhtoons, Balochis and migrants from India and Afghanistan are living here. A study was therefore, undertaken to analyze DNA sequence of the 362 nucleotides of pre S gene of randomly selected HBV isolates from patients with hepatitis B in Karachi. The objective of the study, was to determine the genotypes and subgenotypes and to analyze the spectra of mutations through phylogenetic analysis of HBV strains prevalent in Pakistan.

Results

Prevalence of HBV/D as the most common genotype

Genotype D (219/315) was found to be the most prevalent (70%) amongst our patients. The rest of the genotypes A and a mixture of A and D (AD) were distributed as 20%, and 10% respectively (Table 1).

Dominance of subgenotype D1 and D3

The sequence of nucleotide and the amino acid residue are shown in Figures 1 and 2, respectively. Among the samples studied through phylogenetic analysis (Figure 3), all 13 belonged to genotype D, including 11 to subgenotype D1 and 2 to D3.

Frequency and Distribution of Mutations in the PreS Gene

The nucleotide sequences of the samples with genotype D were compared to those of reference D1 sequences [AY741798](#) from Iran, [AY161159](#) from India, [AY796032](#) from Turkey and reference D3 sequences [AY373430](#) from India, [EI00615](#) from Eastern India, [X85254](#) from Italy and [V01460](#) from France. In the 11 samples of genotype D, mutations were detected at 27 different sites and insertion of nucleotide C was noted after nt 2893 in sample 562HBV2 (Figure 1). Amino acid sequences of part of the pre S protein (14–135) for the HBV isolates from the study samples with genotype D were compared with the amino acid sequences of the same reference sequences (Figure 2).

Discussion

On the basis of 362 nucleotides sequence in the pre S gene, this study demonstrates that Hepatitis B virus geno-

type D with subgenotype D1 and D3 is circulating in Karachi, Pakistan. Previously, several studies have been conducted which report Pakistan among many regions with intermediate to high HBV sero-prevalence, with D being the most prevalent genotype [4,5]. Our study was based on patients who, either were settlers in Karachi but originally belonged to different provinces or had come to Karachi for treatment. The pattern of genotype prevalence found through this study is in line with the studies from South East Asia, especially the countries sharing borders with Pakistan and thus confirms the high proportion of genotype D in southeastern Asia.

In Asia, research on genotyping, was initially conducted extensively in Japan and China, therefore, the genotypes of these countries namely B & C were considered as the most prevalent genotypes of Asia. Later on, when research was extended to other countries, predominance of D was found in South Asia and the Middle East such as India, Afghanistan and Iran [6-9]. The genotype D, in particular, happens to be most widely distributed and found universally, with its highest prevalence in a belt stretching from Southern Europe and North Africa to India, and in West and South Africa, having intravenous drug users as the most affected population within these countries [10,11]. This study, therefore, highlights the genotypic link between various ethnic groups within the country and people of the neighboring countries.

Phylogenetic analysis of genotype D could distinguish four subgenotypes, D1 to D5 [1,12]. The geographical distribution of the subgenotypes within D was found less restricted than that of genotypes A, B, and C, although the strains from Middle East mainly belonged to D1, those from South Africa and Alaska to D3, while those from Oceania and Somalia to D4. Whereas, D5 was only reported from Japan and India. The majority of the strains in this study belonged to subgenotype D1(84.6%). This subgenotype has mainly been shown in strains from Turkey, India, Iraq, Iran and Israel [11]. History of population migration from the Middle East to Pakistan is well known, which explains the predominance of HBV/D1 infection in

Table 1: Genotype distribution in the population studied.

Genotype	Male (n = 237)		Female (n = 78)		Total (n = 315)	
	Number	Percent	Number	Percent	Number	Percent
A	45	19	20	25	65	20
D	170	72	49	63	219	70
AD	22	9	9	12	31	10

3187

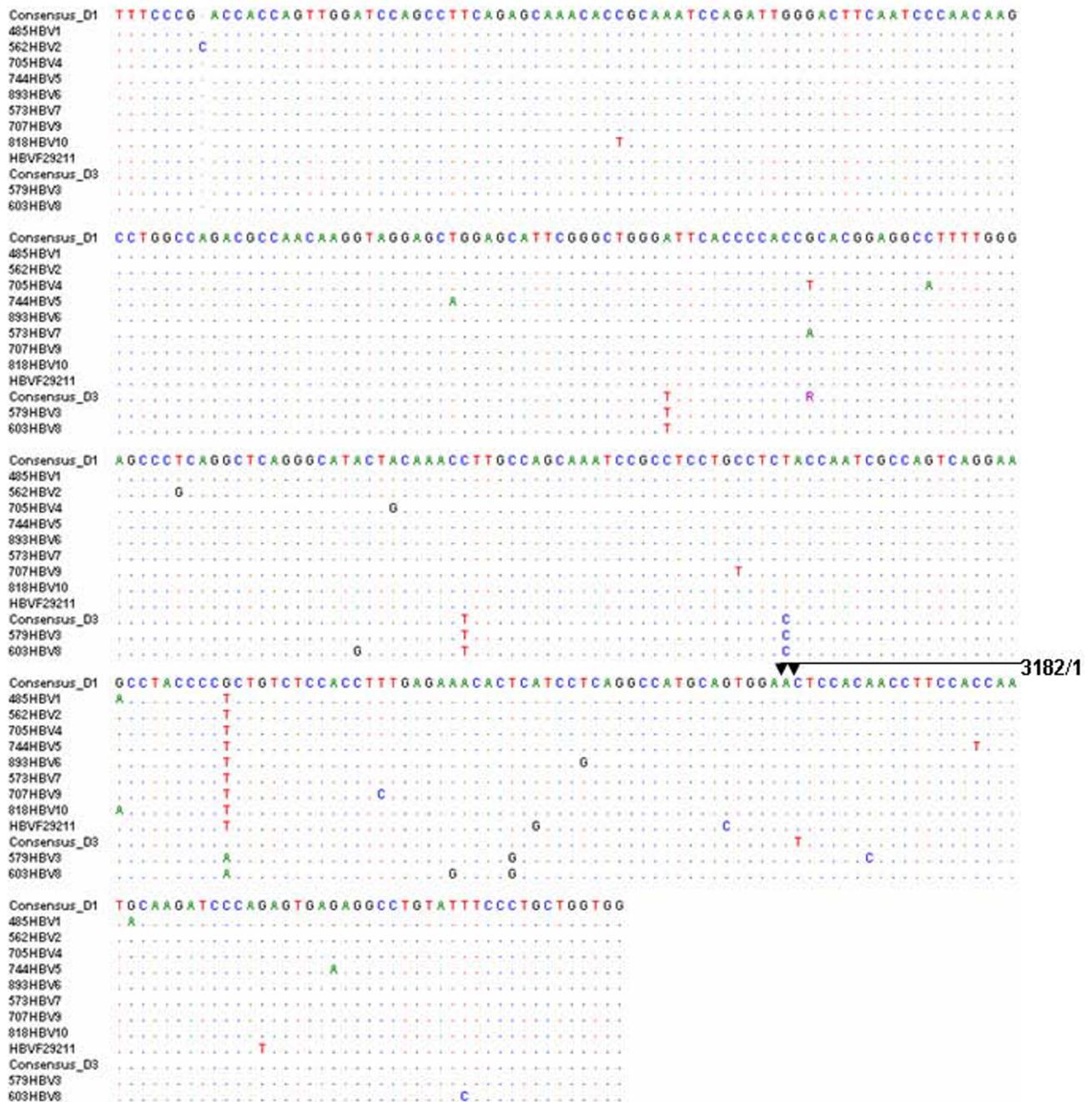


Figure 1
 The nucleotide sequences of the samples with genotype D were compared to those of reference D1 sequences AY741798 from Iran, AY161159 from India, AY796032 from Turkey and reference D3 sequences AY373430 from India, E100615 from Eastern India, X85254 from Italy, V01460 from France.

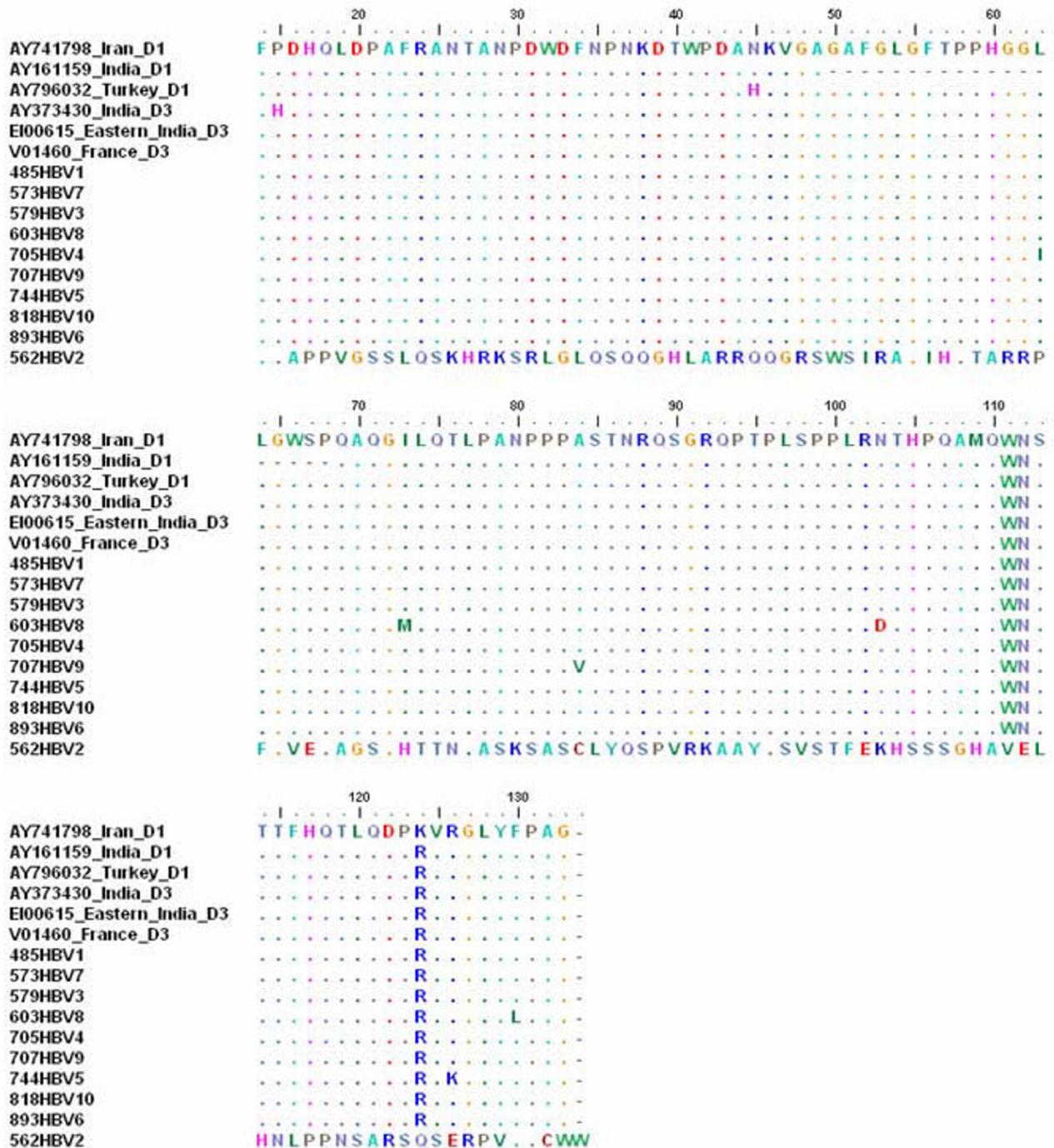


Figure 2
 Amino acid sequences (14–135) of part of the pre S protein for the HBV isolates from the study samples compared with the amino acid sequences of the same reference sequences.

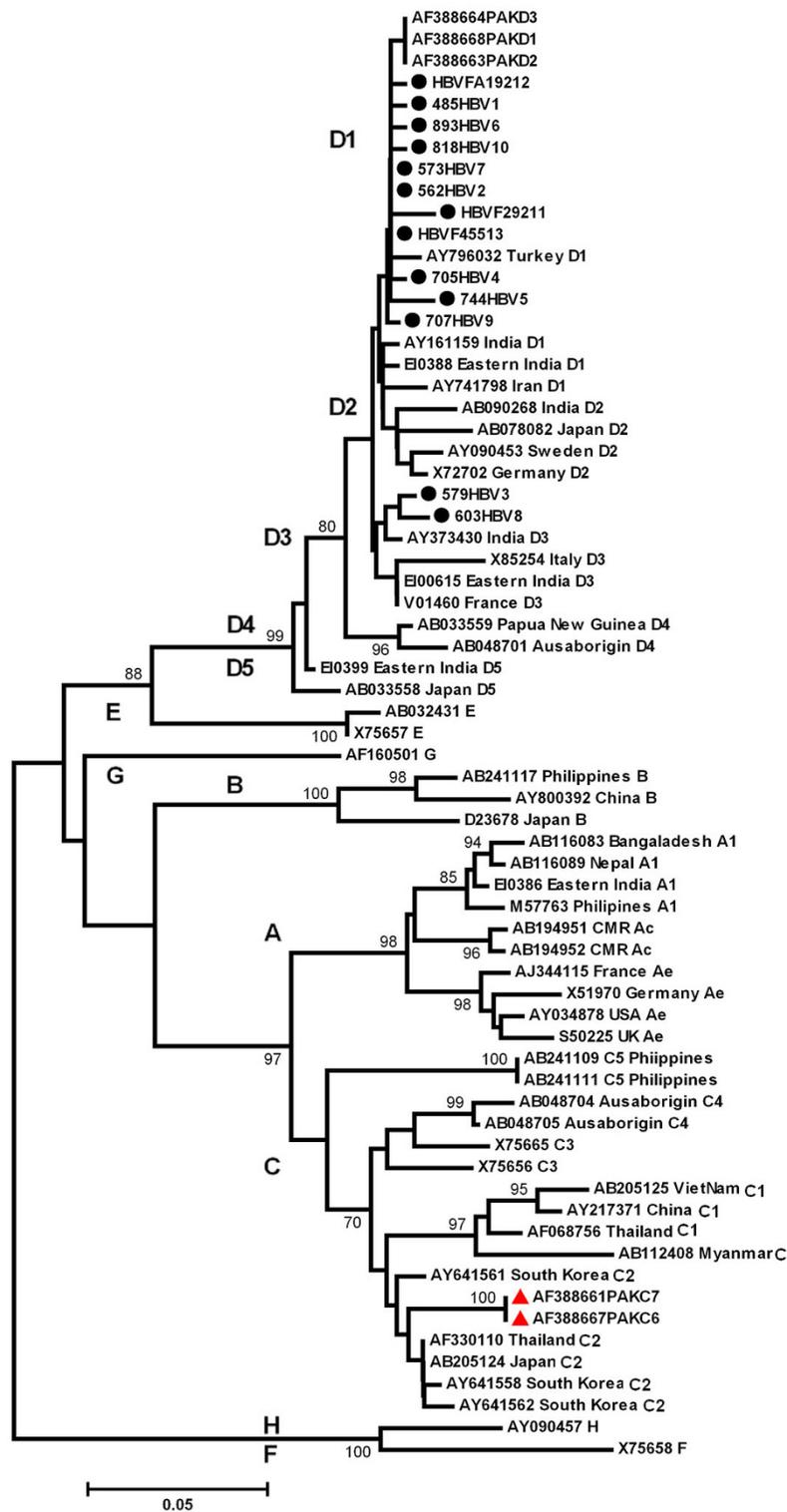


Figure 3
Phylogenetic tree obtained by the neighbor-joining method. Strains from this study are matched with strains obtained from GenBank indicated with their accession numbers. The designation and origin of the strains are given at the nodes of the branches.

this population. In neighboring India, genotype D is predominant with presence of subgenotypes D1, D2, D3 and D5 [1,12]. None of these subgenotypes was characterized by specific amino acid substitutions [11].

Two strains which clustered on a branch within the D3 samples [EF432767](#) and [EF584672](#) were more similar to each other than to the other D3 strains and seem to be derived from a phylogenetically related common ancestor. The *ayw3* specifying strains within this subgenotype D3 have so far been reported from IDUs in Europe and the USA [11]. More studies are needed to investigate if the D3 strains have had fewer introductions into Pakistan compared to the D1 strains. The isolates studied, were found closely related to each other rather than with other D sequences, clustering near each other having 98.1–99% identity between the sequences. This close relationship demonstrates a consistent circulation of closely related viruses within this relatively multidimensional population.

Besides genotype D (76.3%), being the most prevalent, 20% of patients of the present study were found to be infected with genotype A, whereas 10% were co-infected with both types A and D. A recombination between genotypes A and D has also been reported from India [13]. Presence of mixed genotypes A and D in this study indicated a need for further studies for evaluation of possible recombination in samples from Pakistan. Studies from India have generally reported a similar pattern, predominance of genotype D coexisting with varying proportion of genotype A. HBV genotype D has also been described as the genotype of intravenous illicit drug users [14]. In a Pakistani study on IDUs by Alam et al. [5], the genotype distribution was found to be quite similar to the one in our study; genotype D at 62.5%; genotype A, 8.9% while 28.5% individuals were found to be infected with a mixture of genotype A & D. In our patients, the A and D combination was 10% compared to 28.5% reported earlier in Pakistani IDUs. This three times high percentage explains the cross contamination due to sharing of the syringes. Even triple infection has been reported in 0.9% of HBV infected intravenous drug users from areas where 3 HBV genotypes A, B and C are prevalent such as in China [15].

Conclusion

This study confirms the predominance of genotype D in southeastern Asia. The sub-genotypes D1 and D3, dominate in the Pakistani patients. All isolates closely matched with those reported from Iran, India and Egypt. More studies are required to find out the genotypic heterogeneity as a result of intermixing with strains due to population migration.

Methods

Specimens

Blood was drawn from 350 patients undergoing treatment at Pakistan Medical Research Council (PMRC) and Ziauddin

University Hospital (ZUH), diagnosed positive for HBV by HBsAg (ELISA – MUREX kit by Abbott Laboratories). The patients were randomly selected irrespective of their age and gender and a written consent (parental consent in case of under 18 years of age) was obtained. The study was approved by the Ziauddin University ethical review committee.

DNA extraction and Amplification by PCR

DNA was extracted from 200 µl of serum using extraction kit (QIA AMP DNA mini kit 250 reactions Cat #51306). A modified version of nested PCR developed by Naito et al. [16] was followed for amplification. The step one PCR was carried out with universal outer primers P1 (sense) and S1-2 (antisense) in a tube containing 50 µl of a reaction buffer made up of the following components: 50 ng of each outer primer, a 200 µM concentration of each of the four deoxynucleotides, 1 U of Taq polymerase (Perkin-Elmer, Norwalk, Conn.), and 10× PCR buffer containing 1.5 mM MgCl₂. The thermocycler (GeneAmp PCR system 2400, 9600, and 9700; Perkin-Elmer) was programmed to first incubate the samples for 2 min at 94°C, followed by 35 cycles consisting of 94°C for 20 s, 55°C for 20 s, and 72°C for 1 min with a final extension of 5 min at 72°C and incubation at 4°C. Two second-round PCRs were performed for each sample, with the common universal sense primer B2 used as the inner primer (sense) with a combination called mix A for genotypes A, B and C. Mix B consisted of antisense primer B2R with a combination of Mix B for genotypes D, E and F. A 2.5 µl aliquot of the first PCR product was added to two tubes containing the second sets of each of the inner primer pairs, each of the deoxynucleotides, Taq polymerase and PCR buffer, as in the first reaction and amplified for 35 cycles with: preheating at 94°C for 3 min, 15 cycles of amplification at 94°C for 20 s, 58°C for 30 s, and 72°C for 40 s, and an additional 20 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 45 s with an extension of 7 min. at 72°C and incubation at 4°C. Genotypes of HBV for each sample were determined by identifying the genotype-specific DNA bands on electrophoresis. Positive PCR results were repeated twice for confirmation.

Sequencing

The PCR product of 1065 bp obtained with the universal primers P1 and S1-2 was sequenced in 13 HBV genotype D strains on an automated DNA sequencer by Gene Link USA.

Data analysis

Phylogenetic analysis was done using sequences from various countries retrieved from the international DNA database (DDBJ/EMBL/GenBank).

The best and the high scoring matches with our sequences were aligned with the CLUSTAL W software program [17] and the alignment was confirmed by visual inspection.

Phylogenetic analysis of a 362 nt fragment of the S gene employing MEGA, version 2.0 [18], and PHYLIP, version 3.54c [19], was the basis for HBV genotyping using Kimura's 2-parameter algorithms with the neighbour-joining method. The reliability of different phylogenetic groupings was evaluated using the bootstrap test (1000 bootstrap replications). As almost identical groupings were observed with these tree-building programs, only the MEGA-based tree is presented.

The nucleotide sequence data reported in this paper has been submitted to the DDBJ/EMBL/GenBank databases and can be retrieved with accession number [EF432765](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) through [EF432777](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) at <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>, EMBL in Europe and the DNA Data Bank of Japan.

ExpAsy Translate tool, retrieved at <http://us.expasy.org/> was used to translate the nucleotides (DNA) sequence of s protein to amino acid sequence. The deduced amino acid residues were sequentially set up in Figure 2.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SB and AAS designed the Research project. SB carried out the molecular genetic studies, the sequence alignments and drafted the manuscript. TM and AAS helped in the bench work and optimization of the procedure. RC assisted in the Phylogenetic analysis of the sequences and helped to draft the manuscript. SB and RC wrote the final manuscript.

Acknowledgements

The authors wish to thank Dr. Sirajudaullah and 'The Lab' for their help in collection of HBV samples. We also wish to thank and acknowledge Dr. Ali Akhtar for sequencing of the isolates. Thanks are also due to Mr Mumtaz Baig who spared his time to go through numerous manuscript revisions and Ms. Tayyab Unnisa for assistance in bench work.

References

- Norder H, Couroucé AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH, Locarnini S, Magnus LO: **Genetic diversity of hepatitis B virus strains derived worldwide: Genotypes, subgenotypes, and HBsAg subtypes.** *Intervirology* 2004, **47**:289-309.
- Norder H, Couroucé AM, Magnus LO: **Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes.** *Virology* 1994, **198**:489-503.
- Kramvis A, Weitzmann L, Owiredu WK, Kew MC: **Analysis of the complete genome of subgroup A' hepatitis B virus isolates from South Africa.** *J Gen Virol* 2002, **83**:835-839.
- Abbas Z, Muzaffar R, Siddiqui A, Naqvi SAA, Rizvi SAH: **Genetic variability in the precore and core promoter regions of hepatitis B virus strains in Karachi.** *BMC Gastroenterology* 2006, **6**:20.
- Alam MM, Zaidi SZ, Shaikat S, Sharif S, Angez M, Naeem A, Saleha S, Butt JA, Malik SA: **Common Genotypes of Hepatitis B virus prevalent in Injecting drug abusers (addicts) of North West Frontier Province of Pakistan.** *Virology J* 2007, **4**:63.
- Toan NL, Song LH, Kreamsner PG, Duy DN, Binh VQ, Koeberlein B, Kaiser S, Kandolf R, Torresi J, Bock CT: **Impact of Hepatitis B virus genotypes and genotype mixtures on the course of liver disease in Vietnam.** *Hepatology* 2006, **25**:1375-1384.
- Thakur V, Guptan RC, Kazim SN, Malhotra V, Sarin SK: **Profile, spectrum and significance of HBV genotypes in chronic liver disease patients in the Indian subcontinent.** *J Gastroenterol Hepatol* 2002, **17**:165-170.
- Amini BOS, Alavian SM, Adeli A, Sarrami-Forooshani R, Sabahi F: **Hepatitis B virus genotyping, core promoter and precore/core mutations among afghan patients infected with hepatitis B: a preliminary report.** *J Med Virol* 2006, **78**:358-364.
- Amini BOS, Sarrami-Forooshani R, Mahboudi F, Sabahi F, Adeli A, Noorinayer B, Azizi M, Zali M: **Genotype characterization and phylogenetic analysis of hepatitis B virus isolates from Iranian patients.** *J Med Virol* 2005, **75**:227-34.
- Flodgren E, Bengtsson S, Knutsson M, Strebkova EA, Kidd AH, Alexeyev OA, Kidd-Ljunggren K: **Recent high incidence of fulminant hepatitis in Samara, Russia: molecular analysis of prevailing hepatitis B and D virus strains.** *J Clin Microbiol* 2000, **38**(9):3311-3316.
- Bläckberg J, Braconier JH, Widell A, Kidd-Ljunggren K: **Long-term outcome of acute hepatitis B and C in an outbreak of hepatitis in 1969-72.** *Eur J Clin Microbiol Infect Dis* 2000, **19**:21-26.
- Banerjee A, Kurvanob F, Datta S, Chandra PK, Tanaka Y, Mizokami M, Bhattacharya SK, Chakravarty R: **Phylogenetic relatedness and genetic diversity of HBV genotype strain isolated from Eastern India.** *J Med Virol* 2006, **78**:1164-1174.
- Datta S, Banerjee A, Chandra PK, Mahapatra PK, Chakrabarti S, Chakravarty R: **Drug trafficking routes and hepatitis B in injection drug users, Manipur, India.** *Emerg Infect Dis* 2006, **12**(12):1954-1957.
- van Ameijden EJ, van den Hoek JA, Mientjes GH, Coutinho RA: **A longitudinal study on the incidence and transmission patterns of HIV, HBV and HCV infection among drug users in Amsterdam.** *Eur J Epidemiol* 1993, **9**:255-262.
- Chen BF, Kao JH, Liu CJ, Chen DS, Chen PJ: **Genotypic dominance and novel recombinations in HBV genotype B and C co-infected intravenous drug users.** *J Med Virol* 2004, **73**:13-22.
- Naito H, Hayashi S, Abe K: **Rapid and Specific Genotyping System for Hepatitis B Virus Corresponding to Six Major Genotypes by PCR Using Type-Specific Primers.** *J Clin Microbiol* 2001, **39**:362-374.
- Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.
- 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.
- Felsenstein J: *PHYLIP inference package version 3.5* Department of Genetics, University of Washington, Seattle; 1993.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

