Author/s:
Thurnheer, MC;Edwards, R;Schulz, TR;Yuen, L;Littlejohn, M;Revill, P;Bannister, E;Chu, M;Tanyeri, F;Wade, A;Biggs, BA;Sasadeusz, J

Title:
Genotypic profiles of hepatitis B in African immigrants and their clinical relevance

Date:
2017-06-01

Citation:
Thurnheer, M. C., Edwards, R., Schulz, T. R., Yuen, L., Littlejohn, M., Revill, P., Bannister, E., Chu, M., Tanyeri, F., Wade, A., Biggs, B. A. & Sasadeusz, J. (2017). Genotypic profiles of hepatitis B in African immigrants and their clinical relevance. Journal of Medical Virology, 89 (6), pp.1000-1007. https://doi.org/10.1002/jmv.24732.

Research Article

Genotypic profiles of Hepatitis B in African immigrants and their clinical relevance†

M Christine Thurnheer[1], Rosalind Edwards[3], Thomas R Schulz[1,2], Lilly Yuen[3], Margaret Littlejohn[3], Peter Revill[3], Liz Bannister[3], Melissa Chu[4], Firuz Tanyeri[4,] Amanda Wade[5], Beverley-Ann Biggs[1,2], Sasadeusz Joseph[1*]

[1] Victorian Infectious Diseases Service, The Royal Melbourne Hospital, at the Doherty Institute, 792 Elizabeth Street, Melbourne Victoria 3000
[2] University of Melbourne, Department of Medicine/ RMH, at the Doherty Institute, 792 Elizabeth Street, Melbourne Victoria 3000
[3] Victorian Infectious Diseases Reference Laboratory, at the Doherty Institute, 792 Elizabeth Street, 792 Elizabeth Street , Melbourne Victoria 3000
[4] University of Melbourne, Department of Medicine, Royal Parade, Parkville Victoria, 3050
[5] Department of Infectious Diseases, University Hospital, Geelong, Ryrie St, Geelong Victoria 3220

***Corresponding author**
Dr Joe Sasadeusz
Victorian Infectious Diseases Service, The Royal Melbourne Hospital, at the Doherty Institute, 792 Elizabeth Street, Melbourne Victoria 3000
Phone: 613 9342 7212 Email: J.Sasadeusz@mh.org.au

Shortened title: Hepatitis B Genotypes in African Immigrants

†This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/jmv.24732]

Additional Supporting Information may be found in the online version of this article.

Received 21 July 2016; Revised 24 October 2016; Accepted 9 November 2016
Journal of Medical Virology

DOI 10.1002/jmv.24732

ABSTRACT

Hepatitis B virus (HBV) from 40 adult African immigrants in Australia was characterised to determine the prevalence of different HBV genotypes and subgenotypes. A mutational analysis was then performed to determine the presence of clinically significant mutations and correlate them to clinical outcomes. Initial sequencing analysis revealed 13 with genotype A (32.5%), 13 with genotype D (32.5%), and 14 with genotype E (35%). Serology showed that 37 were HBeAg negative. Phylogenetic analysis identified a high prevalence (25%) of HBV subgenotype A1 in our cohort, a subgenotype which has been associated with more aggressive clinical disease.

BCP/PC sequencing was obtained for 38 patients. BCP and/or PC mutations were identified in 36/38 (95%). The median viral load of all patients was 2995 IU/mL and most of the pathology results were within the normal range. Only one patient had an increased APRI score of 1.1 suggestive of cirrhosis.

We present novel information on the HBV genotypes amongst the African population in Australia along with clinical correlates. The high prevalence of A1 subgenotype in this population supports the current Australian recommendation to commence hepatocellular carcinoma screening in Africans with chronic HBV from 20 years old.

KEY WORDS

Chronic hepatitis B; Australia; Serology; Mutation; Sequencing; Hepatocellular carcinoma

INTRODUCTION

Chronic Hepatitis B virus (HBV) infection affects approximately 248 million people worldwide (24). Hepatitis B virus prevalence is highest in Africa and eastern Asia (24) and chronic infection will lead to death in up to 25% of those infected due to hepatocellular carcinoma (HCC) and liver cirrhosis (14, 15).

In Africa some 75 million people have chronic HBV (14, 22) where it is considered hyper endemic, with a seroprevalence of >8% (12). In this setting the virus is predominantly transmitted either vertically through mother to child transmission or horizontally during early childhood (8). Current evidence regarding the geographic distribution of HBV genotypes includes that genotype D is found in north Africa, genotype E is found in western and central regions and genotype A is in eastern and southern regions (12). Genotypes E and A1 are unique to Africa, except in cases where there is migration history such as slave trade or refugees.

In Australia, an estimated 218,000 people (approximately 1% of the total population) are living with chronic HBV, the majority (56%) of who were born overseas (20). (Asia/Pacific: 38%, Europe: 10%, Africa/Middle East: 7%, Americas: 0.8%). In the past two decades there have been increasing numbers of refugees arriving from sub-Saharan Africa to Australia and also to other parts of the world (2, 21). People from culturally and linguistically diverse (CALD) backgrounds (particularly Asia-Pacific or Sub-Saharan African background) are listed as one of the priority populations of the Australian National Hepatitis B Strategy (7) and as such are an important group to consider and prioritise.

Preliminary work from an ongoing study of HBV in African children (4) in Australia documented the presence of genotypes A, D and E in this population group. Phylogenetic analysis revealed the presence of sub genotype A1 which is associated with aggressive clinical disease (9), as well as a high prevalence of HBV variants encoding mutations in the basal core promoter (BCP) and/or precore (PC) gene associated with hepatitis B e antigen (HBeAg) negativity, with the majority of subjects having sero-converted to anti-HBe (HBeAb) despite their young age.(4) The

identification of BCP variants in particular is important, as recent studies have shown the presence of these mutations is associated with increased likelihood of progression to advanced liver disease, particularly cirrhosis (23, 25)

In this study we aimed to characterise HBV isolated from adult African immigrants and refugees living in Victoria, Australia by determining the prevalence of different HBV genotypes/sub genotypes, and then performing a mutational analysis to determine the presence of clinically significant mutations and correlating them to clinical outcomes.

MATERIALS AND METHODS

Ethics approval was received from the Royal Melbourne Hospital human research and ethics committee.

Patient Samples

Stored serum samples were obtained from 42 African patients from three separate sites in Victoria – University Hospital, (Geelong), Isis Primary Care (Hoppers Crossing) and the Royal Melbourne Hospital hepatitis and travel/refugee clinics.

Clinical data

Sociodemographic and clinical data were collected retrospectively from medical charts together with pathology results, and were entered into a standardised case report form (CRF). The sociodemographic parameters collected were date of birth, sex, ethnicity, country of origin and age. Data were censored on 30th December 2013. Regions of origin were defined according to the geographic location of the country of origin. Pathology results collected included aspartate aminotransferase (AST), alanine amino transferase (ALT), bilirubin, albumin, gamma glutamyltransferase (GGT), alkaline phosphatase (ALP), platelets, HBeAg status and HBV viral load (IU/mL). AST to platelet ratio index (APRI) scores were calculated as a marker of hepatic fibrosis where concurrent results for platelet counts and AST were available (29/40 patients), using

the following formula: ((AST (IU/L) /AST upper limit of normal (ULN) (IU/L))/platelet Count (10^9/L))*100. (40IU/L was used as ULN for AST values).

The presence or absence of clinical signs of cirrhosis (including ascites, encephalopathy and splenomegaly) were recorded where available. Information on clinical management (current treatment yes/no, drug used for treatment; imaging studies including transient elastography results) was recorded where available.

Clinical information was compiled in a separate database for further analysis. Statistical analysis was performed using STATA SE12© (StataCorp, College Station, Texas, USA). Fisher's exact test was used to compare categorical data of multiple groups. Kruskal Wallis test was used for numerical data comparison, assuming non-normal distribution of the data.

HBV DNA Extraction, PCR and Sequencing

HBV DNA was extracted from 200µL serum using the QIAamp DNA Minikit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Extracted DNA was eluted in a final volume of 50µL of supplied elution buffer. To determine genotype/subgenotype, a 893bp fragment of the reverse transcriptase (RT) region of the polymerase gene (and corresponding overlapping region of the surface (S) gene) was amplified and sequenced using primers Seq2 and 2996 (Table 1). To determine the presence of BCP and PC mutations, amplification and sequencing of a 356bp fragment of the BCP/PC region was performed using primers PC5 and PC2 (Table 1).

Amplification was carried out using HotStarTaq *Plus* DNA Polymerase (QIAGEN, Hilden, Germany) at a concentration of 2 units/reaction, 10 x PCR Buffer (containing 15mM MgCl2), 0.16µM dNTPs and 0.2µM of each primer. For the BCP/PC PCR, the final concentration of $MgCl_2$ was slightly higher at 2mM. Amplification conditions were as follows: denaturation at 95°C for 5 min, followed by 50 cycles of denaturation (94°C for 30 s), annealing (55°C for 30 s) and extension (72°C for 40 s), then a final extension step of 72°C for 10 min. Prior to sequencing, amplified

products were purified using ExoSAP-IT (USB® Products Affymetrix, Inc.) according to the manufacturer's instructions.

Sequencing was carried out using the Big Dye® Terminator v3.1 Cycle Sequencing Kit (Applied Bio systems, Foster City, CA), according to the manufacturer's instructions. Full genome amplification was carried out as previously described (6) using HotStarTaq *Plus* DNA Polymerase (QIAGEN, Hilden, Germany) and primers P1 and P2 (Table 1). After purification, sequencing was performed using nine overlapping primers – SeqP1, 903, 865, 1798, 2254, JM, 2996, OSX1 and SeqP2 (Table 1).

HBV consensus sequences for the polymerase and BCP/PC genes, as well as the full genomes, were constructed using the DNA sequence analysis program *SeqScape* (ABI Prism, Applied Bio systems, Foster City, CA). HBV genotype and unique HBV mutations were identified using a web-based analysis program, *SeqHepB,* as previously described (5, 27)

Phylogenetic Analysis

The HBV polymerase sequences obtained were aligned using Clustal W and Bio edit (www.mbio.ncsu.edu/BioEdit/bioedit.html). Consensus sequences were compared to a set of published reference sequences (GenBank) representing the 10 human HBV genotypes (A-J, approx. 50 sub genotypes) (Figure 1). Neighbour-joining phylogenetic trees were constructed using the MEGA software, version 5.0 (Tamura et al., 2011) and evaluated by bootstrap resampling with 1000 replicates.

RESULTS

Study population, clinical characteristics

Of the 42 patients recruited, two had insufficient viral load for amplification and subsequent genotyping and these two patients were not included in the analysis.

The clinical characteristics of the 40 patients included in the study are presented in Table 2. Most patients were male (28, 70%), the median age was 36 years (IQR 30-45years), and the majority (26, 65%) originated from East Africa or Horn of Africa (14 from Sudan, 6 from Somalia, 4 from Ethiopia, 2 from Eritrea, Table 2). Five originated from West Africa (1 from Ghana, 3 from Liberia, 1 from Nigeria) and five were from the central and southern regions of Africa (1 from Burundi, 1 from Zimbabwe and 3 from DR Congo). No information on the country of origin was available for four patients. According to clinical classifications none of the patients had stigmata suggestive of chronic liver disease.

The median viral load of all patients was 2995 IU/mL (IQR 583 – 34333 IU/mL). Most of the pathology results were within the normal range, and APRI score results for most of the 29 patients with available results for both platelet counts and AST lay below the threshold for increased risk for severe fibrosis or cirrhosis (Table 3). One patient had an increased APRI score of 1.1 suggestive of cirrhosis (> 1.0 indicative of cirrhosis)(16). He was a 26 year old man from Liberia, infected with HBV Genotype E, with a high viral load of 9,254,000 IU/mL, increased AST and ALT levels (79 IU/L and 89 IU/L respectively) and a normal platelet count of $184x10^9$ /L. He was HBeAb positive and HBeAg negative. No results for transient elastography were available for this patient.

Only seven patients were receiving antiviral treatment (Table 3). Entecavir was the most commonly used drug (5/7 patients), one individual received tenofovir and one person had a history of treatment with Pegylated Interferon (Peg IFN). The median viral load measured whilst on treatment was 470 IU/mL (IQR 263-1331 IU/mL).

Results for transient elastography were available for 6/40 patients, the mean result for median stiffness was 4.3 kPa. One patient had a high result of 15.5 kPa, suggestive of advanced fibrosis. He was a 37 years old man from Zimbabwe, infected with HBV genotype E. His viral load was 469,000 IU/ml. Pathology results revealed increased hepatic enzymes (ALT 107, AST 66, Bilirubin 11, GGT 107 and ALP 123). He was negative for HBeAg and positive for HBeAb. No APRI score

was calculated for this patient as synchronous pre-treatment AST and platelet results were not available and he was commenced on treatment after the sample collection.

Genotyping

Initial sequencing analysis of the 40 patients revealed 13 with genotype A (32.5%), 13 with genotype D (32.5%), and 14 with genotype E (35%). Recent reclassification of genotype A sub genotypes (11) has grouped A3, A4 and A5 together as quasi-subgenotype A3, and A6 reclassified as A4. In our cohort, based on sequencing of the polymerase region, subtyping of genotype A viruses revealed 10 patients with subgenotype A1 (25%) 2 patients (5%) with quasi-subgenotype A3 and one with subgenotype A4 (2.5%). (Figure 1)

A recent reclassification of sub genotypes of genotype D, concluded that there are six (D1-D6) sub genotypes, not eight as previously thought (26). The low inter-subgenotype divergence between D3 and D6 suggests that they belong to one subgenotype. In addition, subgenotype D7 (from Tunisia, Algeria and Morocco) then becomes subgenotype D6, and subgenotype D8 (from Niger) (11) is considered a genotype D/E recombinant not a separate subgenotype. In our cohort, based on sequencing of the polymerase region, subtyping of genotype D isolates revealed five patients (12.5%) with subgenotype D2 and eight patients (20%) with either subgenotype D6 or D/E recombinants (Figure 1). These recombinant viruses have also been identified in a cohort of African children living in Australia.(4)

Full genome amplification and sequencing of two of these isolates revealed clustering with isolates from Niger (1), thus confirming the presence of D/E recombinant viruses. (Phylogenetic analysis for full genome sequencing not shown) Full genome analysis could not be obtained for the remaining 6 isolates due to low viral load, therefore we can only classify them as subgenotype D6, clustering with isolates from Tunisia.

Geographic distribution

The geographic distribution of the isolated genotypes is shown in Table 2 and Figure 2. Twenty-six patients originated from the east African/Horn of Africa region, and these patients were infected with either genotype E (8), D6 (5), D2 (5), D/E recombinant (2) or A1 (6). The five patients from West Africa were infected with genotype E (3) and quasi-subgenotype A3 (2). No genotype A1 was isolated from this region. The five individuals from central and southern Africa were infected with genotype A1 (3), A4 or E (one each). Comparison of genotype distribution showed no statistically significant variation among the different regions, most likely attributable to the limited number of patients.

Mutations associated with antiviral resistance or basal core promoter/precore regions

Six patients were on antiviral therapy at the time of testing (five on entecavir, one on tenofovir) and one patient had received PegIFN for 11 months in 2011. None of these patients had viruses with mutations associated with antiviral resistance in their pre-treatment samples. No sequencing of on-treatment samples was performed. One patient had an rtA181T mutation in the polymerase (associated with resistance to lamivudine, telbivudine and adefovir) in the absence of antiviral therapy. The same patient had an sP120T mutation in the surface which is associated with HBV immunoglobulin or vaccine escape mutants.

BCP/PC sequencing was obtained for 38 patients. BCP and/or PC mutations were identified in 36/38 (95%) (Figure 3). Two patients did not have the classic BCP or PC mutations; however, one patient had a 1 nucleotide (nt) insertion in PC at amino acid (aa) 5 whilst the other had a 1nt deletion in PC at aa12.

HBe serology

Serology results are shown in Table 3: Only a minority of 3/40 patients were HBeAg positive while 37 were HBeAg negative. A high number of 34 /37 patients with no detectable HBeAg were HBeAb positive (HBe seroconverted), for the remaining three HBeAg negative

patients no results for HBeAb were available. The three patients with positive HBeAg and negative HBeAb results had mutations in the BCP region only.

DISCUSSION

Hepatitis B is thought to have been present in Africa for thousands of years and possibly much longer (17). The predominant genotype is thought to be A1 with probably a more recent establishment of Genotype E (3) Genotype D is also prevalent in Northern Africa and the Middle East. (28)

We have described the prevalence of various genotypes of HBV amongst a population of African immigrants to Australia. Our data demonstrated a wide variety of genotypes with a particularly high prevalence of genotype A1 (25%) and genotype E (35%). Genotypes tended to cluster in specific geographical areas with Genotype A1 predominating in the Horn of Africa and central Africa while Genotype E predominating in the Horn of Africa. There was however significant overlap of genotypes in geographical regions.

Phylogenetic analysis identified a high prevalence (25%) of HBV subgenotype A1 in our cohort. In young males A1 has been associated with more aggressive clinical disease, with a 4.5 times higher relative risk of developing hepatocellular carcinoma (HCC) compared to non-A genotypes (12). Subgenotype A1 HBV encodes mutations in the BCP and PC region that decrease core protein translation (10) and reduce HBeAg expression, (13) although the mechanism for the rapid onset of HCC in subjects infected with this strain remains unclear. It was notable that no HCC occurred amongst this small sample, the majority of whom had mild liver disease and were HBeAg negative, although longitudinal follow up of these patients is required to monitor their cancer risk.

The current recommendations for HCC screening in patients with Hepatitis B suggest Africans over 20 years old should have regular screening and this finding supports this

recommendation (19). The high prevalence of genotype A, which is known to be responsive to interferon, may suggest this as a suitable therapeutic option in this population. The use of interferon in Genotype A infection can potentially remove the need for lifelong nucleoside/tide therapy, especially in young individuals of child-bearing age.

Although the high rate of HBeAg negativity observed in an African paediatric HBV population (4) suggested the possibility of more advanced liver disease in this cohort, particularly as they entered adulthood, our data demonstrated mild liver disease in the vast majority of individuals with only one case of documented cirrhosis and one of possible cirrhosis. It was notable that both these cases were Genotype E.

Consistent with this high rate of HBeAg negativity, a high incidence of BCP and/or PC mutations were observed amongst all genotypes. Mutations in these regions have been associated with increased progression to cirrhosis and/or HCC, (18, 25) again supporting guidelines for early age of onset for screening in Africans with chronic HBV. The high rates of these mutations also raises the possibility that routine use of genotyping may assist in the selection of patients for treatment, as has been recommended for Asian patients with genotype B or C virus (23).

Limitations to this study include that the sample size was small and the numbers of each genotype did not allow for accurate between groups statistical comparisons. Data was also collected retrospectively, some information was incomplete and the majority of the patients were from countries in the horn of Africa.

In conclusion, these findings have important implications for patient monitoring and treatment, particularly with increasing immigration from the Sub-Saharan region. They also highlight the need for further study into the disease course of the various genotypes of HBV, particularly those with high prevalence in African populations.

ACKNOWLEDGEMENTS

We would like to thank Ms Christalla Hajisava for her assistance with the submission.

REFERENCES

1. Abdou Chekaraou MS, et al. 2010. A novel hepatitis B virus (HBV) subgenotype D (D8) strain, resulting from recombination between genotypes D and E, is circulating in Niger along with HBV/E strains. J Gen Virol 91:1609-20.
2. ABS 2015, posting date. Australian Bureau of Statistics. [Online.]
3. Andernach IE, et al. 2009. Slave trade and hepatitis B virus genotypes and subgenotypes in Haiti and Africa. Emerg Infect Dis 15:1222-8.
4. Bannister, E. G., L. Yuen, M. Littlejohn, R. Edwards, S. Locarnini, W. Hardikar, and P. A. Revill. 2012. Presented at the Hepatology.
5. Bartholomeusz A S, et al. 2004. Molecular modelling of hepatitis B virus polymerase and adefovir resistance identifies three clusters of mutations. Hepatology 40:246A.
6. Gunther SB, et al. 1995. A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients. J Virol 69:5437-44.
7. Department of Health, Australian Government. 2014. Second National Hepatitis B Strategy 2014-2017. http://www.health.gov.au/internet/main/publishing.nsf/Content/C353814FE3255962CA257BF0001DE841/$File/Hep-B-Strategy2014-v3.pdf
8. Hoffmann CJ and Thio CL. 2007. Clinical implications of HIV and hepatitis B co-infection in Asia and Africa. Lancet Infect Dis 7:402-9.
9. Kew MCA, et al. 2005. Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu-speaking sub-saharan Africans. J Med Virol 75:513-21.
10. Kimbi G CM, et al. 2012. The effect of the G1888A mutation of subgenotype A1 of hepatitis B virus on the translation of the core protein. Virus Res 163:334-40.

11. Kramvis, A. 2014. Genotypes and genetic variability of hepatitis B virus. Intervirology 57:141-50.

12. Kramvis A and Kew MC. 2007. Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. Hepatol Res 37:S9-S19.

13. Kramvis A and Kew MC. 2007. Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. Hepatol Res 37:S27-S32.

14. Lavanchy D. 2004. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. J Viral Hepat 11:97-107.

15. Liaw YF and Chu CM. 2009. Hepatitis B virus infection. Lancet 373:582-92.

16. Lin Z H, et al. 2011. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: an updated meta-analysis. Hepatology 53:726-36.

17. Littlejohn MS, et al. 2016. Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus. Cold Spring Harb Perspect Med 6(1):a021360.

18. Liu CJ, et al. 2006. Role of hepatitis B virus precore/core promoter mutations and serum viral load on noncirrhotic hepatocellular carcinoma: a case-control study. J Infect Dis 194:594-9.

19. Lok AS and McMahon BJ. 2009. Chronic hepatitis B: update 2009. Hepatology 50:661-2.

20. MacLachlan J H, et al. 2013. The burden of chronic hepatitis B virus infection in Australia, 2011. Aust N Z J Public Health 37:416-22.

21. OECD World Migration in Figures. 2013 [Online.] http://www.oecd.org/els/mig/World-Migration-in-Figures.pdf

22. Ott JJ, et al. 2012. Global epidemiology of hepatitis B virus infection: new estimates of age-specific HBsAg seroprevalence and endemicity. Vaccine 30:2212-9.

23. Revill P and Locarnini S. 2014. Viral factors and predicting disease outcomes in chronic hepatitis B. Gut. 64:191-3.

24. Schweitzer AJ, et al. 2015. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. Lancet 386(10003):1546-55

25. Tseng TC, et al. 2015. Higher proportion of viral basal core promoter mutant increases the risk of liver cirrhosis in hepatitis B carriers. Gut 64:292-302.

26. Yousif M and Kramvis A. 2013. Genotype D of hepatitis B virus and its subgenotypes: An update. Hepatol Res 43:355-64.

27. Yuen LKW, Ayres A, Littlejohn M, Colledge D, Edgely A, Maskill WJ, Locarnini SA, and Bartholomeusz A. 2007. SEQHEPB: A sequence analysis program and relational database system for chronic hepatitis B. Antiviral Research In Press, Uncorrected Proof.

28. Zehender GE, et al. Enigmatic origin of hepatitis B virus: an ancient travelling companion or a recent encounter? World J Gastroenterol 20:7622-34.

Table 1: HBV primers used for PCR and/or sequencing[a]

Name	Position	Sequence (5'-3')
Seq2	300-318	TTG GCC AAA ATT CGC AGT C
2996	1175-1193	GCG TCA GCA AAC ACT TGG C
PC5	1604-1623	TCG CAT GGA GAC CAC CGT GA
PC2	1940-1959	GGC AAA AAC GAG AGT AAC TC
P1	1821-1841	*****CCG GAA AGC TTG AGC TCT TCT TTT TCA CCT CTG **CCT AAT CA**
P2	1806-1825	*****CCG GAA AGC TTG AGC TCT TCA AAA AGT TGC ATG **GTG CTG G**
SeqP1	1821-1841	TTT TTC ACC TCT GCC TAA TCA
SeqP2	1806-1825	AAA AAG TTG CAT GGT GCT GG
865	2301-2320	TTY GGA GTG TGG ATT CGC AC
903	2337-2362	GTT GAT AAG ATA GGG GCA TTT GGT GG
1798	3237-3257	CCA CTG CAT GGC CTG AGG ATG
2254	412-433	GAA GAT GAG GCA TAG CAG CAG G
JM	3114-3134	TTG GGG TGG AGC CCT CAG GCT
OSX1	807-827	TTT TCT TTT GTC TTT GGG TAT
OS1	2846-2868	GCC TCA TTT TGT GGG TCA CCA TA

[a]Numbering from *Eco*R1 site

*****non-HBV sequence underlined

Table 2: Clinical Data

	All	Viral sub genotypes							p[b)]
HBV (sub)genotype	n (%)	A1	A4	Quasi A3	D2	D6	D/E recomb[a)]	E	
Total	40	10	1	2	5	6	2	14	
Male sex	28 (70)	6	1	1	4	4	1	11	0.845*
Age years (Median)	36	39	37	27	31	41	42	37	0.462**
Region of origin									
East Africa/Horn of Africa	26 (65)	6	0	0	5	5	2	8	0.108*
Eritrea	2								
Somalia	6								
Ethiopia	4								
Sudan	14								
West Africa	5 (12.5)	0	0	2	0	0	0	3	
Ghana	1								
Liberia	3								
Nigeria	1								
Central and Southern Africa	5 (12.5)	3	1	0	0	0	0	1	
Burundi	1								
Zimbabwe	1								
DR Congo	3								
unknown	4 (10)	1	0	0	0	1	0	2	

a) recomb = recombinant

b) comparing genotypes

*Fisher's exact test (categorical data, small sample size) ** Kruskal Wallis (numerical data, non-normal distribution)

Table 3: Pathology Results

(Median values with interquartile ranges (IQR)

Pathology results	***All (IQR)***	***A1***	***A4***	***Quasi A3***	***D2***	***D6***	***D/E***	***E***	**P [a]**
Viral load (U/mL)	3143 (652 - 34333)	2117	2084154	122810	1448	534	8778974	3680	0.152*
ALT (U/L)	26 (18 - 40)	22	37	54	22	21	38	33	0.784*
AST (U/L)	31 (24 -37)	36	36	37	30	25	31	30	0.838*
GGT (U/L)	24 (19 - 52)	20	26	44	25	23	30	36	0.310*
Bilirubin (µmol/L)	10 (7 -12)	11	9	10	7	9	8	11	0.220*
ALP (U/L)	78 (62 - 100)	71	107	81	64	76	79	86	0.639*
Albumin (g/L)	39 (36-41)	37	39	42	41	38	38	39	0.216*
Platelet count (10^9/L)	195 (165 - 228)	161	156	203	215	206	214	200	0.446*
APRI score (29/40 patients)	0.4 (0.3 - 0.5)	0.5	0.6	0.4	0.3	0.5	0.3	0.3	0.304**
Transient elastography	7	2	0	0	2	0	1	2	
Median stiffness (kPa)	4.3 (3.9 - 6)	4.5			2.6		4.3	10.5 (max 15)	
Treatment yes	7	2	0	0	1	1	1	2	0.880*

Tenofovir	1	1							
Entecavir	5	1				1	1	2	
PegINF	1				1				
Viral load on treatment	470 (263-1331)								
Serology									
HBeAg pos	3	1	1	0	1	0	0	0	0.059*
HBeAg neg	37	9	0	2	4	6	2	14	
HBeAb pos	34	9	0	2	4	5	2	12	0.319*
HBeAb neg	3	1	1	0	1	0	0	0	
HBe seroconverted	34	9	0	2	4	5	2	12	0.530*

a) comparing genotypes

*Fisher's exact test (categorical data, small sample size) ** Kruskal wallis (numerical data, non-normal distribution)

All values are medians unless stated, IQR is the interquartile range.

FIGURE LEGEND

Figure 1: Phylogenetic trees of a 761bp fragment of the HBV surface gene, including 143 reference sequences (Accession numbers given in supplementary table). Sequences isolated from this study identified by black square and country of origin. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (>50% shown). (A) Genotype A sequences. (note: quasi-A3 was previously sub-genotypes A3, A4 and A5. A4 was previously sub-genotype A6, see text for explanation.) (B) Genotype D sequences. (note: D3 was previously sub-genotypes D3 and D6. D6 was previously sub-genotype D7.) (C) Genotype E sequences. (D) Phylogenetic tree of HBV full genome sequences, constructed by the neighbour-joining method, showing the D/E recombinant sequences clustering with the isolates from Niger (21).

Figure 2: Hepatitis B genotype/subgenotype distribution by African country of origin.

Figure 3: Graph of the number of patients where HBV BCP/PC mutations were detected according to genotype. Note that multiple mutations were detected in 16 patients.

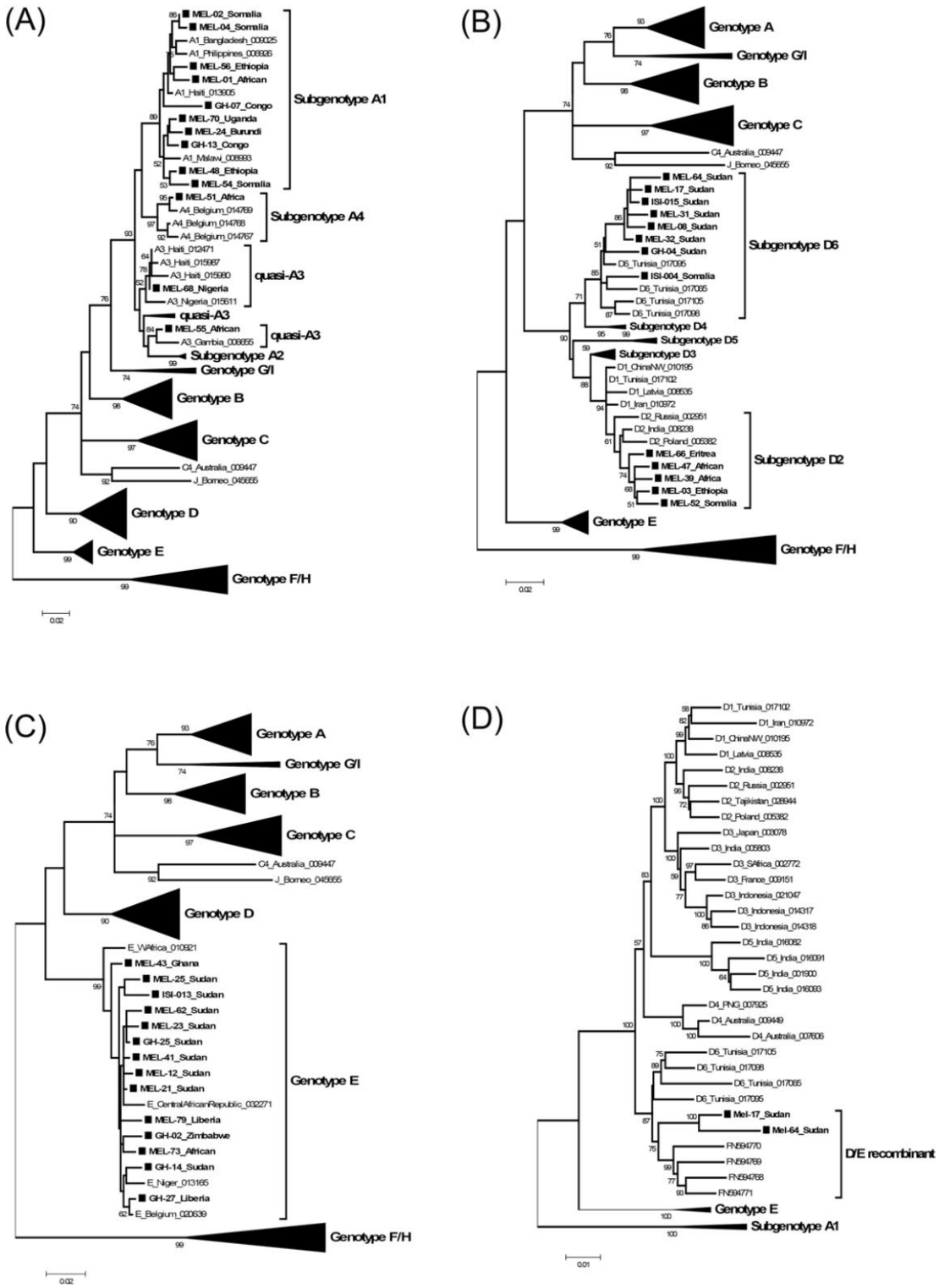

Figure 1

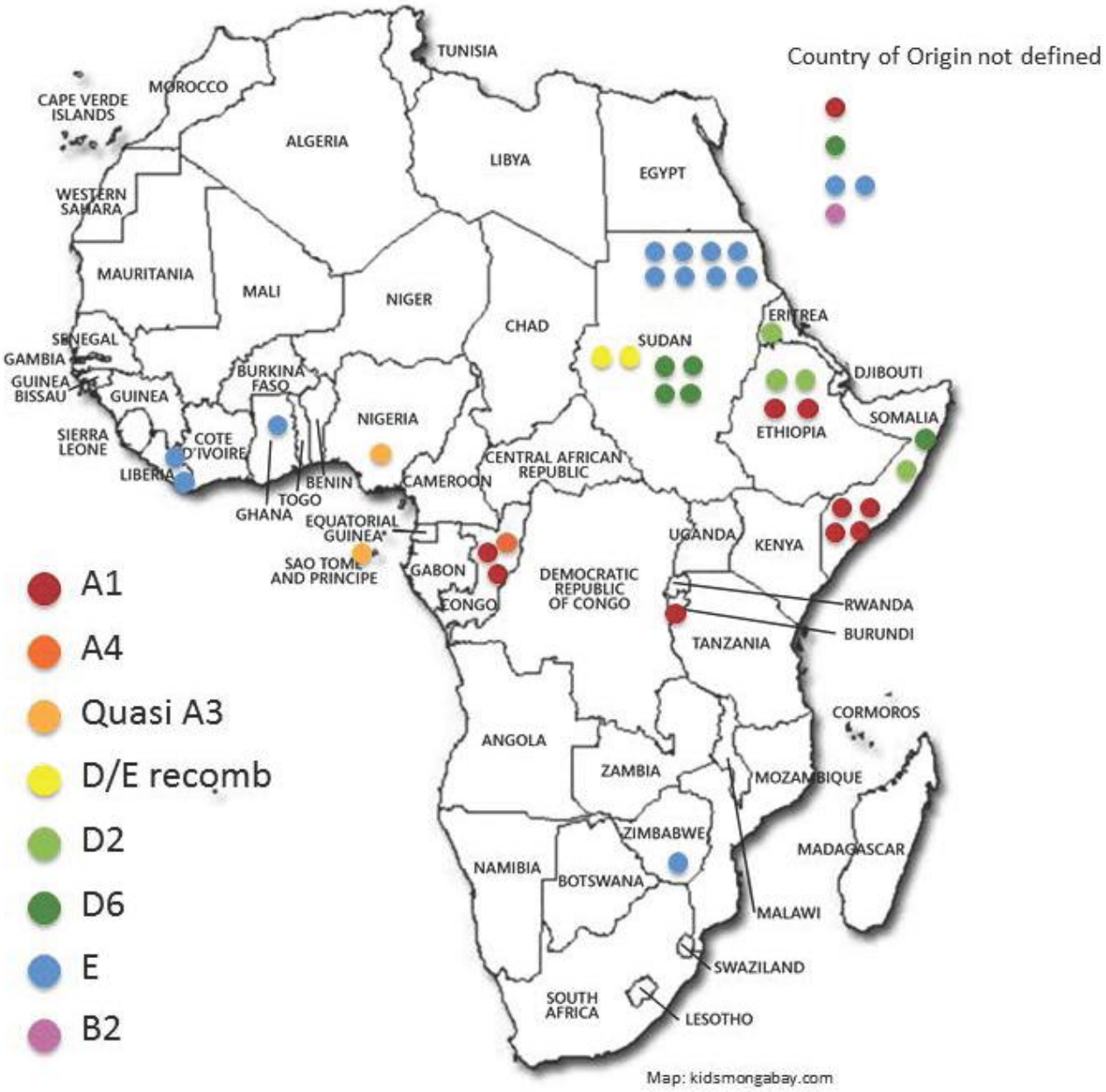

Figure 2

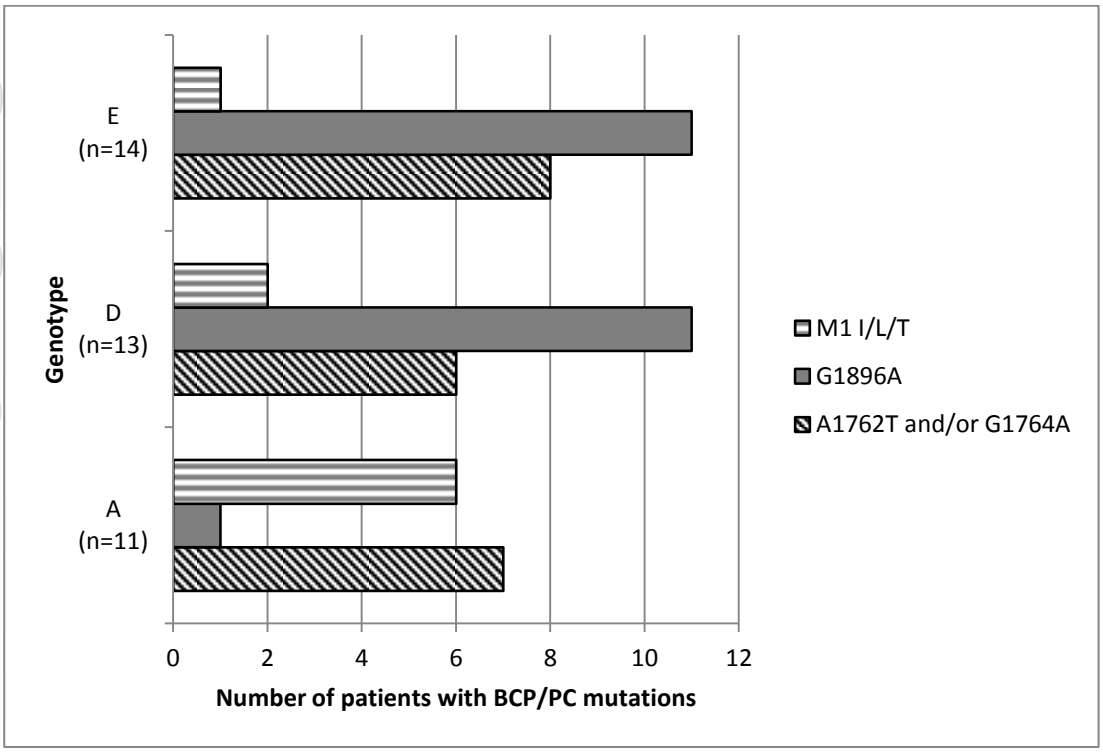

Figure 3

