



Brief report

A putative hepatitis B virus sequence motif associated with hepatocellular carcinoma in South African adults



Tongai G. Maponga^{a,b,†}, Anna L. McNaughton^{c,†}, Cori Campbell^c, Mariateresa de Cesare^c, Jolynne Mokaya^d, Sheila F. Lumley^{c,e}, David Bonsall^{c,e,f}, Camilla LC Ip^c, Haiting Chai^c, Christo Van Rensburg^g, Richard H. Glashoff^{b,h}, Elizabeth Waddiloveⁱ, Wolfgang Preiser^{a,b}, Jason T. Blackard^j, M. Azim Ansari^c, Anna Kramvis^k, Monique I. Andersson^{a,c,e}, Philippa C Matthews^{i,l,m,*}

^a Division of Medical Virology, Stellenbosch University, Faculty of Medicine and Health Sciences, Tygerberg, Cape Town, South Africa

^b National Health Laboratory Service, Tygerberg Business Unit, Tygerberg, Cape Town, South Africa

^c Nuffield Department of Medicine, University of Oxford, Oxford, UK

^d Parasites and Microbes, Wellcome Sanger Institute, Hinxton, UK

^e Department of Infectious Diseases and Microbiology, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK

^f Big Data Institute, Old Road, Oxford OX3 7FZ, UK

^g Division of Gastroenterology, Stellenbosch University, Faculty of Medicine and Health Sciences, Tygerberg, Cape Town, South Africa

^h Immunology Unit, Division of Medical Microbiology, Stellenbosch University, Faculty of Medicine and Health Sciences, Tygerberg, Cape Town, South Africa

ⁱ The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

^j University of Cincinnati College of Medicine, Cincinnati, OH, USA

^k Hepatitis Virus Diversity Research Unit, Department of Internal Medicine, University of the Witwatersrand, 7 York Road, Parktown, 2193 Johannesburg, South Africa

^l Division of Infection and Immunity, University College London, Gower Street, London WC1E 6BT, UK

^m Department of Infectious Diseases, University College London Hospital, Euston Road, London NW1 2BU, UK

ARTICLE INFO

Article History:

Received 7 June 2024

Accepted 14 November 2024

Available online 20 February 2025

Keywords:

Hepatitis B virus

Hepatocellular carcinoma

Liver cancer

Africa

Whole genome sequencing

ABSTRACT

Introduction and Objectives: Chronic hepatitis B virus (HBV) infection is a major risk factor for hepatocellular carcinoma (HCC). In African populations, HCC frequently presents at an advanced stage with poor outcomes. We applied whole genome sequencing (WGS) to compare HBV genomes in individuals with and without HCC.

Materials and Methods: We identified adults with HBV infection, with and without complicating HCC, in Cape Town, South Africa. We generated HBV WGS using pan-genotypic probe-based enrichment followed by Illumina sequencing.

Results: Compared to the non-HCC group, HCC patients were more likely to be male ($p < 0.0001$), older ($p = 0.01$), HIV-negative ($p = 0.006$), and have higher HBV viral loads ($p < 0.0001$). Among 19 HCC and 12 non-HCC patients for whom WGS was obtained, genotype A dominated (74 %), of which 96 % were subgenotype A1. PreS2 deletions ($\Delta 38-55$) were enriched in HBV sequences from HCC patients ($n = 7$). The sequence motif most strongly associated with HCC comprised either a deletion or polymorphism at site T53 in PreS2 – collectively coined ‘non-T53’ – together with a basal core promoter (BCP) mutation G1764A (AUROC = 0.79).

Conclusions: In this setting, HBV sequence polymorphisms and deletions are associated with HCC, and ‘non-T53 + G1764A’ represents a putative signature motif for HCC. Additional investigations are needed to disaggregate the impact of other demographic, clinical, and environmental influences, to ascertain the extent to which viral polymorphisms contribute to oncogenesis, and to determine whether HBV sequence is a useful biomarker for HCC risk stratification.

Abbreviations: AUROC, Area Under the Receiver Operating Curve; BCP, Basal Core Promoter; BEAST, Bayesian Evolutionary Analysis by Sampling Trees; bp, Base Pairs; DNA, Deoxyribonucleic Acid; ER, Endoplasmic Reticulum; Genbank, Genetic Sequence Database; HIV, Human Immunodeficiency Virus; HBV, Hepatitis B Virus; HCC, Hepatocellular Carcinoma; IU/mL, International Units per Milliliter; MCMC, Markov Chain Monte Carlo; nt, Nucleotide; PRJEB, Project Identifier in European Nucleotide Archive (ENA); STATA, Statistical Software for Data Science; VL, Viral Load; WHO, World Health Organization; WGS, Whole Genome Sequencing; X gene, Hepatitis B Virus X gene

* Corresponding author.

E-mail address: philippa.matthews@crick.ac.uk (P.C. Matthews).

† These authors contributed equally.

<https://doi.org/10.1016/j.aohep.2024.101763>

1665-2681/© 2025 Fundación Clínica Médica Sur, A.C. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

In the World Health Organization (WHO) Africa region, an estimated 65 million individuals are living with chronic hepatitis B virus (HBV) infection [1,2]. The high incidence of hepatocellular carcinoma (HCC) in Africa (63,000 cases per annum in 2018) [3,4] reflects the prevalence and distribution of HBV infection [5,6]. HBV-associated HCC in African populations affects adults in early and middle life, who typically present with advanced malignancy, leading to high mortality, with substantial individual and societal impact [7,8]. International targets that aim to eliminate the public health threat of HBV infection include HCC as an important area of focus [9].

Multiple HBV genotypes circulate in Africa, including genotypes A, D and E, with genotypes A and E potentially being of particular significance in contributing to HCC incidence [10]. In Southern Africa, infection with subgenotype A1 HBV has a strong association with HCC [11,12]. Malignant transformation can occur as a result of chronic inflammatory/fibrotic liver disease, HBV DNA integration into the host genome, cell stress caused by the accumulation of aberrant viral proteins [13], and/or a direct influence of HBV genes (particularly the X gene) [14]. Specific viral polymorphisms have been associated with HCC, including truncated genes, pre-core insertions/deletions ('indels'), and basal core promoter (BCP) mutations [15–17]. Such sequence changes can be used potentially to infer disease risk or prognosis [18–20]. However, further work is needed to better describe the mutational landscape of HBV in diverse viral genotypes, advance insights into specific associations between viral sequence polymorphisms and HCC, and determine their mechanistic impact. Evaluation is needed to determine whether viral sequence can inform clinical risk assessment, surveillance or interventions.

We performed HBV sequence analysis from South African adults to explore viral sequence polymorphisms in those with and without HCC.

2. Material and Methods

2.1. Study samples

We retrospectively drew on banked serum samples from adults with a confirmed diagnosis of chronic HBV infection, in cohorts with and without HCC; the clinical and epidemiological features of these have been previously published (see [21] and [22], respectively) (Suppl methods 1; Suppl Fig. 1).

2.2. Illumina sequencing, genome assembly, and phylogenetic analysis

We generated HBV sequence data based on an adapted version of a published Illumina protocol [23] (Suppl methods 2). HBV reads were mapped to reference sequences (genotypes A–I) prior to generating consensus WGS. A standard HBV reference strain (Genbank accession X02763, genotype A) was used for numbering positions in the genome.

For phylogenetic analysis, we aligned nucleotide alignments using new sequence data from this cohort, 61 full-length South African sequences from Genbank, and genotype reference sequences [24] using Clustal X 2.1 [25]. We performed phylogenetic inference using a Bayesian Markov Chain Monte Carlo (MCMC) approach as implemented in the Bayesian Evolutionary Analysis by Sampling Trees (BEAST) version 1.10.4 program (Suppl methods 3).

2.3. Analysis

We identified HBV polymorphisms that have been previously associated with HCC [14,26–29] (Table 1), and explored the association between polymorphisms and HCC using area under the receiver-operating curve (AUROC). The frequency of deletions at each site was compared between HCC and non-HCC sequences. Statistical analyses were performed using GraphPad Prism v10 and STATA v17.0.

Table 1
Frequency of HBV polymorphisms previously associated with hepatocellular carcinoma (HCC) identified in consensus sequences derived by Illumina among adults with and without a diagnosis of HCC. Frequencies are based on analysis of consensus sequences. Partial genomes were derived in some cases; hence, the denominator at certain sites is lower than the total number of samples sequenced.

HBV genomic region and amino acid site	Frequency of HBV polymorphism in HCC (n = 19)	Frequency of HBV polymorphism in non-HCC (n = 12)	p-value
HBsAg polymorphisms and deletions [30,40,49,50]			
W172*	0/19	0/12	n/a
W182*	0/19	0/12	n/a
L216*	1/19	0/11	0.63
Pre-S2 T53C (F22 L)	8/18	3/11	0.30
PreS2Δ38–55*	7/17	0/11	0.02
'Non-T53' (Pre-S2 T53C/A/G OR Δ38–55 ^b)	15/19	4/11	0.03
Other pre-S1 deletion	0/17	0/11	n/a
Other S deletion	0/19	0/11	n/a
Pre-core region			
Q2*	1/19	0/12	0.61
S13T	12/19	8/12	0.58
W28*	5/19	2/12	0.44
(also reported as G1896A)			
G29D	3/19	1/12	0.49
(also reported as G1899A)			
BCP region [14], overlapping with X protein changes			
T1753C (I127T in X protein)	4/19	1/11	0.38
A1762T (K130 M in X protein)	12/18	4/11	0.11
G1764A (V131I in X protein)	14/18	5/11	0.09
C1766T (F132Y/I/R ^c in X protein)	6/18	1/11	0.15
T1768A	4/19	1/11	0.38
Double mutation A1762T/G1764A	11/18	4/11	0.18
Triple mutation T1753C / A1762T / G1764A	3/18	0/11	0.22
Deletion spanning 1762–1764	1/19	0/11	0.63

Stop mutation indicated by *.

^a The presence of any deletion spanning the region preS2Δ38–55 was recorded [40].

^b The preS2Δ38–55 deletion and the T53C mutation spanned the same site so were analysed together and considered as 'non T53'.

^c F132Y was the only variant observed at site 132 in our sequences.

2.4. Ethical statement

Ethical approval was granted from the health research ethics committee at the University of Stellenbosch (S13/04/072 and N11/09/284).

3. Results

3.1. HBV-associated HCC with male sex and higher HBV viral loads

We identified samples from 161 adults with chronic HBV infection of whom 68 had HCC and 93 not. Compared to the non-HCC group, those with HCC were more likely to be male (81 % vs. 46 %, respectively, $p < 0.0001$), older (median 41 vs. 36 years, $p = 0.01$), and had higher HBV DNA viral loads (median 5.2 vs 3.5 log₁₀ IU/mL, $p < 0.0001$). HIV coinfection was present in 18 of 65 HCC cases (3 had no documented test result) and 46 of 93 non-HCC cases (27.8 % vs 49.5 %; $p = 0.006$) (Suppl Table 1).

We were able to undertake WGS HBV sequencing from 19 samples from the HCC group and 14 from the non-HCC group (Suppl Methods 2), with sufficient reads to construct WGS assemblies from 31 of these (19 HCC and 12 non-HCC (Table 1; Suppl Fig. 2)). Sequence data can be accessed in GenBank project PRJEB71107.

Overall, genotype A predominated, accounting for 23 of 31 (74.2 %), of which 22 were subgenotype A1. There was no difference in genotype distribution between HCC and non-HCC groups (Suppl Fig. 1; Suppl Tables 2 and 3; Suppl Results 1). HBV sequences from HCC and non-HCC groups were interspersed with other South African HBV sequences, suggesting there was no particular viral lineage associated with the development of HCC (Suppl Fig. 3).

3.2. Deletions and substitutions in HBV Pre-S2 significantly enriched in HCC

We examined HBV sequences for the presence of polymorphisms and deletions previously associated with HCC (Table 1). There was sequence coverage in the PreS2 region in 17 of 19 HCC cases, among which 7 of 17 had PreS2 $\Delta 38-55$ deletions, compared to none in the non-HCC group ($p = 0.02$) (Fig. 1A, B). In the consensus sequences, deletions were observed ranging from 3 bp to 42 bp in length, with a mean of 21 bp. Start locations of the sequences varied, but all deletions terminated by nucleotide (nt) 55. We also evaluated the T53C substitution that has been reported in association with HCC [30]. In the HCC group, wild-type T53 was uncommon, due to a combination of deletions ($n = 6$), and substitutions ($n = 9$; T53C substitution in 8 and T53 G in one). Thus, overall, the 'non-T53' motif occurred in 15/19 samples in the HCC group compared to 4/11 in the non-HCC group ($p = 0.027$; Table 1).

3.3. BCP mutations

Mutations in the BCP region associated with HCC by other studies (A1762T, G1764A, and the combination of both mutations) were more common in the HCC group; however, this was not significantly different from the non-HCC group (Table 1).

3.3.1. Associations of combined polymorphisms with HCC

Finally, we analysed 'non-T53' together with the most common BCP mutation (G1764A) or the BCP double mutation (A1762T + G1764A). The combination of 'non-T53' with the G1764A mutation was the most strongly associated with HCC (AUROC = 0.79) (Table 1, Fig. 1C).

4. Discussion

In this small retrospective sample set, HBV PreS2 'non-T53' (either as a result of a mutation or deletion) in combination with the BCP G1764A polymorphism, is the sequence motif most strongly associated with HCC. Both of these changes have been reported independently in South African HBV sequence data [31–35]. However, to the best of our knowledge, this combined motif has not been previously studied. Further data are needed to evaluate the sensitivity and specificity of the association between this motif and HCC.

In this small cohort, we could not determine the impact of potential confounders such as age, sex, HBV VL, HIV status, and antiviral treatment. Furthermore, there may be an impact of additional contributors which were not measured, such as host genetic/epigenetic factors, family history, coinfection with other blood-borne viruses, metabolic liver disease, obesity, diet, alcohol, and external environmental factors such as toxin exposure [36–39].

PreS2 deletions were observed in HBV sequences from people presenting with HCC in our cohort, and have also been reported in association with HCC in West African sequence data, potentially associated with aflatoxin exposure and/or conferring a viral fitness advantage [40]. Mechanistically, 'non-T53' may have an influence on PreS2 function or regulation and/or impact the spacer domain of the polymerase protein. Proposed oncogenic mechanisms associated with these motifs include the accumulation of defective proteins in the endoplasmic reticulum (ER), stress responses resulting in DNA damage, centrosome over-duplication, and genomic instability [28,41–43]. PreS2 deletions start at a wide range of sites (sometimes prior to nt 38) but consistently terminate immediately prior to nt 55, where there is a highly conserved region of approximately 20 bp, and the first of several cysteine residues within a putative zinc finger domain, considered to be essential for reverse transcriptase activity [44]. Therefore, deletions downstream of nt 55 are likely to be detrimental to viral replication, potentially explaining why they all terminate within the same region.

A longer duration of infection may be needed for cumulative BCP mutations to develop, and an increased frequency of these polymorphisms has been reported as HBV infection progresses [45,46]. The increase in A-T-rich regions in the BCP region (nt 1762–1770) seen in HCC-associated sequences in this study may be associated with upregulating viral transcription [47], and these mutations may influence malignant transformation mediated through the overlapping X gene [19].

In future, larger studies are needed, supported by an unbiased, sequence-agnostic approach to WGS changes (instead of focusing only on polymorphisms that have been previously reported). There is a need to interrogate the impact of HIV coinfection, as this may affect oncogenic risk through its effect on inflammation and immune responses (although, conversely, people living with HIV may derive protection from earlier and more consistent treatment [48]). Longitudinal follow-up in a multi-centre study would be needed to determine whether relevant HBV mutations pre-date HCC, or evolve as cancer develops.

To date, HBV sequencing protocols have been limited by high VL thresholds and/or cost. As sequencing only generates data for VL samples above a certain threshold, there may be bias in our dataset, e.g. with artificial enrichment for A1762T and G1764A mutations (lower VL samples are not represented). Retrieval of whole or partial viral sequences is also influenced by variables including quality and quantity of stored material, extent of host DNA contamination, and potentially viral genotype. As the sensitivity and efficiency of pan-genotypic sequencing methods improves, it may be possible to sequence HBV genomes at lower VL from smaller sample volumes, and to generate long-read sequences to support WGS haplotype analysis [23].

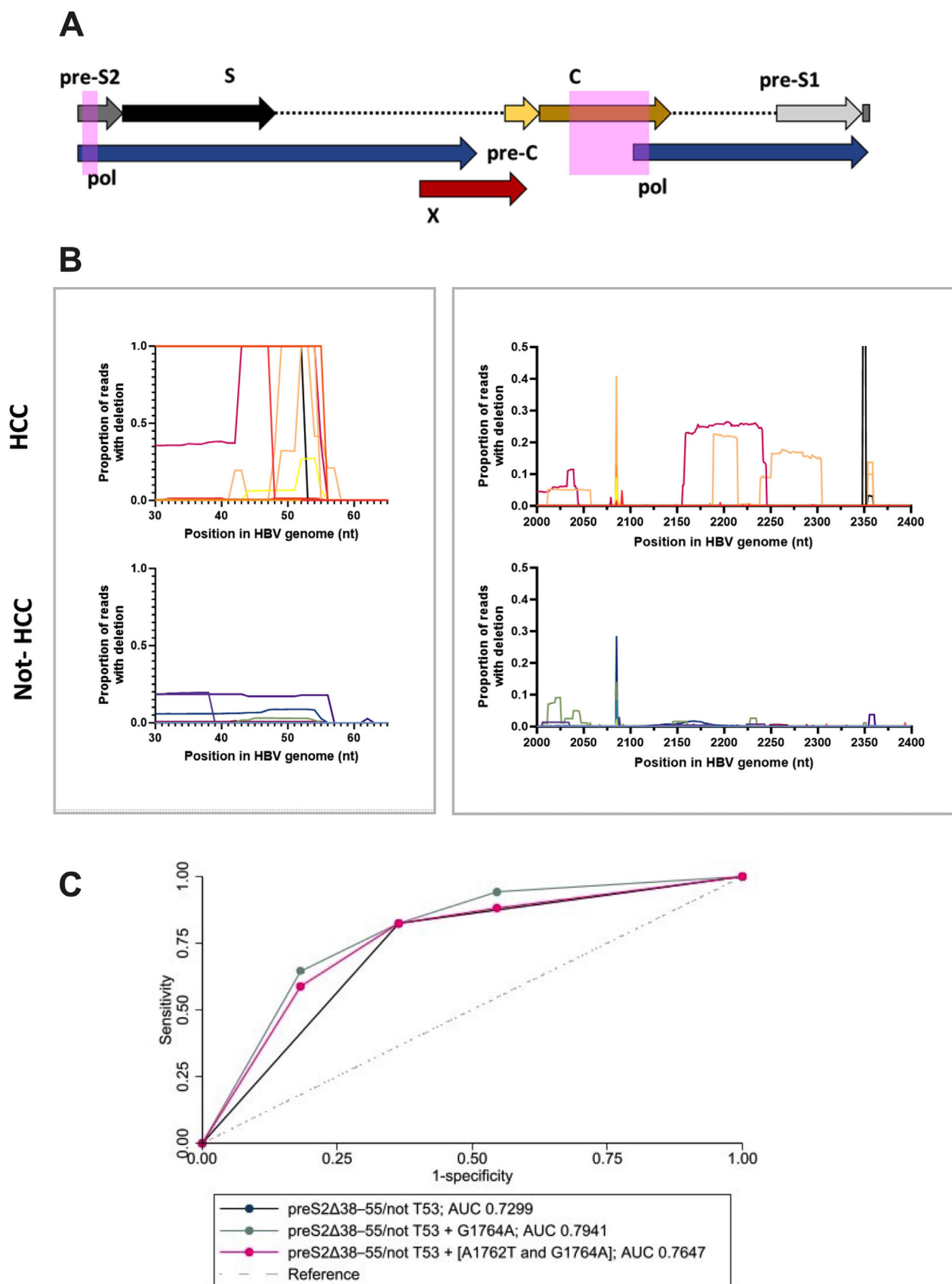


Fig. 1. HBV Sequence motifs associated with HCC in South African adults. A: Schematic of HBV genome. Region of PreS2 and BCP polymorphisms are shown in pink highlights. B: Regions of the HBV genome showing the proportion of reads with deletions at each site. Data are presented for individuals with HCC (top) and without HCC (bottom). Termination of deletions in 7 HBV sequences from the HCC group was at nt 52 ($n = 1$), nt 53 ($n = 2$), nt 54 ($n = 2$), and nt 55 ($n = 2$). Additional discussion of the minority variant deletions is provided in Suppl results 2. The bioinformatic pipeline did not provide deletion frequencies for reads when a deletion was the consensus at that site (the site was absent), so these sites have been assigned a frequency of 100 % to generate these plots. C: Receiver operating characteristic (ROC) curves for pre-S2 deletions and T53 polymorphisms combined with BCP

5. Conclusions

In conclusion, we propose a novel HBV sequence motif that may be associated with HCC, and highlight the pressing need for the field to be expanded by the open-source sharing of HBV sequences together with clinical metadata. Further work is required to determine the cause, effect and chronology of these sequence changes relative to clinical and demographic characteristics, and the evolution of HCC.

Author contributions

Conceptualization: TGM, ALM, PCM. Patient Recruitment: TGM, CVR, MIA. Ethics and Governance: TGM, EW, MIA, RG. Laboratory Work: TGM, ALM, MdC, DB. Data Analysis and Curation: ALM, CLCI, MAA, CC, JM, JTB, DB, HC. Original Manuscript Drafting: ALM, CLCI, AK, PCM. Manuscript Editing and Approval: All authors. Supervision: RG, WP, MIA, MAA, PCM.

Funding

PCM received funding support from the Wellcome Trust (grant ref 110,110/Z/15/Z), UCL NIHR Biomedical Research Centre and the Francis Crick Institute (ref. CC2223). TGM received support from the Polio-myelitis Research Foundation, the Harry Crossley Foundation and Columbia University—South Africa Training Program for Research on AIDS-related Malignancies through the National Cancer Institute, NIH (Grant # 1D43CA153715).

Data and materials

Illumina reads and *de novo* assemblies are available from the European Nucleotide Archive (GenBank project PRJEB71107). Consensus sequences are also available on line at <https://www.doi.org/10.6084/m9.figshare.24874557>.

Declaration of interest

PCM has participated in projects supported by GSK, outside the direct scope of the work presented here. MIA has participated in projects supported by Prenetics, J&J, and Pfizer outside the direct scope of this work. AK is an Associate Editor for the journal *Annals of Hepatology*. All other authors declare no conflicts of interest.

Acknowledgements

We are grateful to the staff and patients at our centres of recruitment for their support and participation.

Supplementary materials

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.aohp.2024.101763](https://doi.org/10.1016/j.aohp.2024.101763).

References

- [1] Sonderup MW, Spearman CW. HBV elimination in Africa-current status and challenges. *Clin Liver Dis* 2024;23:e0166. <https://doi.org/10.1097/CLD.0000000000000166>.
- [2] World Health Organization Global Hepatitis Report 2024: Action for Access in Low- and Middle-Income Countries. <https://www.who.int/publications/i/item/9789240091672>. ISBN: 978-92-4-009167-2.
- [3] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424. <https://doi.org/10.3322/caac.21492>.
- [4] Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* 2019. <https://doi.org/10.1038/s41575-019-0186-y>.
- [5] Mak D, Babb de Villiers C, Chasela C, Urban MI, Kramvis A. Analysis of risk factors associated with hepatocellular carcinoma in black South Africans: 2000–2012. *PLoS One* 2018;13:e0196057. <https://doi.org/10.1371/journal.pone.0196057>.
- [6] Brah S, Moussa S, Inoua A, Alhousseini DM, Daou M, Madougou B, et al. Molecular characterization of hepatitis B virus from chronically-infected patients in Niamey, Niger. *Int J Infect Dis* 2016;45:18–23. <https://doi.org/10.1016/j.ijid.2016.02.009>.
- [7] Lemoine M, Thursz MR. Battlefield against hepatitis B infection and HCC in Africa. *J Hepatol* 2017;66:645–54. <https://doi.org/10.1016/j.jhep.2016.10.013>.
- [8] Yang JD, Mohamed EA, Aziz AOA, Shousha HI, Hashem MB, Nabeel MM, et al. Characteristics, management, and outcomes of patients with hepatocellular carcinoma in Africa: a multicountry observational study from the Africa Liver Cancer Consortium. *Lancet Gastroenterol Hepatol* 2017;2:103–11. [https://doi.org/10.1016/S2468-1253\(16\)30161-3](https://doi.org/10.1016/S2468-1253(16)30161-3).
- [9] Global health sector strategy on viral hepatitis 2016–2021. Towards ending viral hepatitis. In: World Health Organization [Internet]. Jun 2016 [cited 7 Nov 2022]. Available: <https://apps.who.int/iris/bitstream/handle/10665/246177/WHO-HIV-2016.06-eng.pdf?sequence=1&isAllowed=y>
- [10] Kedar Mukthinuthalapati VVP, Sewram V, Ndlovu N, Kimani S, Abdelaziz AO, Chiao EY, et al. Hepatocellular carcinoma in Sub-Saharan Africa. *JCO Glob Oncol* 2021;7:756–66. <https://doi.org/10.1200/JGO.20.00425>.
- [11] Kew MC, Kramvis A, Yu MC, Arakawa K, Hodgkinson J. Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu-speaking sub-saharan Africans. *J Med Virol* 2005;513–21. <https://doi.org/10.1002/jmv.20311>.
- [12] Kramvis A. Molecular characteristics and clinical relevance of African genotypes and subgenotypes of hepatitis B virus. *S Afr Med J* 2018;108:17–21. <https://doi.org/10.7196/SAMJ.2018.v108i8b.13495>.
- [13] Hsieh Y-H, Su I-J, Wang H-C, Chang W-W, Lei H-Y, Lai M-D, et al. Pre-S mutant surface antigens in chronic hepatitis B virus infection induce oxidative stress and DNA damage. *Carcinogenesis* 2004;25:2023–32. <https://doi.org/10.1093/carcin/bgh207>.
- [14] An P, Xu J, Yu Y, Winkler CA. Host and viral genetic variation in HBV-related hepatocellular carcinoma. *Front Genet* 2018;9:261. <https://doi.org/10.3389/fgene.2018.00261>.
- [15] Skelton M, Kew MC, Kramvis A. Distinct mutant hepatitis B virus genomes, with alterations in all four open reading frames, in a single South African hepatocellular carcinoma patient. *Virus Res* 2012;163:59–65.
- [16] Chuon C, Takahashi K, Matsuo J, Katayama K, Yamamoto C, Ko K, et al. High possibility of hepatocarcinogenesis in HBV genotype C1 infected Cambodians is indicated by 340 HBV C1 full-genomes analysis from GenBank. *Sci Rep* 2019;9:12186.
- [17] Raihan R, Akbar SMF, Al Mahtab M, Takahashi K, Masumoto J, Tabassum S, et al. Genomic analysis of Hepatitis B virus and its association with disease manifestations in Bangladesh. *PLoS One* 2019;14:e0218744.
- [18] Ge Z, Tian T, Meng L, Song C, Yu C, Xu X, et al. HBV mutations in EnhII/BCP/PC region contribute to the prognosis of hepatocellular carcinoma. *Cancer Med* 2019. <https://doi.org/10.1002/cam4.2169>.
- [19] Siddiqui ZI, Farooqui SR, Azam SA, Afroz M, Wajid S, Parveen S, et al. A comparative study of hepatitis B virus X protein mutants K130M, V131I and KV130/131MI to investigate their roles in fibrosis, cirrhosis and hepatocellular carcinoma. *J Viral Hepat* 2017;24:1121–31.
- [20] Al-Qahtani AA, MR Al-Anazi, Nazir N, Ghai R, Abdo AA, Sanai FM, et al. Hepatitis B virus (HBV) X gene mutations and their association with liver disease progression in HBV-infected patients. *Oncotarget* 2017;8:105115–25.
- [21] Maponga TG, Glashoff RH, Vermeulen H, Robertson B, Burmeister S, Bernon M, et al. Hepatitis B virus-associated hepatocellular carcinoma in South Africa in the era of HIV. *BMC Gastroenterol* 2020;20:226.
- [22] Maponga TG, Andersson MI, van Rensburg CJ, Arends JE, Taljaard J, Preiser W, et al. HBV and HIV viral load but not microbial translocation or immune activation are associated with liver fibrosis among patients in South Africa. *BMC Infect Dis* 2018;18:214.
- [23] McNaughton AL, Roberts HE, Bonsall D, de Cesare M, Mokaya J, Lumley SF, et al. Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). *Sci Rep* 2019;9:7081.
- [24] McNaughton AL, Revill PA, Littlejohn M, Matthews PC, Ansari MA. Analysis of genomic-length HBV sequences to determine genotype and subgenotype reference sequences. *J Gen Virol* 2020. <https://doi.org/10.1099/jgv.0.001387>.
- [25] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23:2947–8.
- [26] Cohen D, Ghosh S, Shimakawa Y, Ramou N, Garcia PS, Dubois A, et al. Hepatitis B virus preS2Δ38–55 variants: a newly identified risk factor for hepatocellular carcinoma. *JHEP Rep* 2020;2:100144.
- [27] Ge Z, Tian T, Meng L, Song C, Yu C, Xu X, et al. HBV mutations in EnhII/BCP/PC region contribute to the prognosis of hepatocellular carcinoma. *Cancer Med* 2019. <https://doi.org/10.1002/cam4.2169>.
- [28] Ahn SH, Park YK, Park E-S, Kim JH, Kim DH, Lim K-H, et al. The impact of the hepatitis B virus polymerase rtA181T mutation on replication and drug resistance is potentially affected by overlapping changes in surface gene. *J Virol* 2014;88:6805–18.
- [29] Liu WC, Wu IC, Lee YC, Lin CP. Hepatocellular carcinoma-associated single-nucleotide variants and deletions identified by the use of genome-wide high-throughput analysis of hepatitis B virus. *J Pathol* 2017;243(2):176–92.

polymorphisms as a predictor of HCC status in HBV sequences. Curves are shown for combinations of PreS2 'non T53' alone, or combined with the BCP polymorphisms G1764A and the double mutation A1764T/G1764A. Two sequences, HCC-34 and nHCC-12 were excluded from the analysis as there was no sequence for the region under consideration. AUC, area under the curve.

- [30] Hosseini SY, Sanaei N, Fattahi M-R, Malek-Hosseini SA, Sarvari J. Association of HBsAg mutation patterns with hepatitis B infection outcome: asymptomatic carriers versus HCC/cirrhotic patients. *Ann Hepatol* 2019;18:640–5.
- [31] Makondo E, Bell TG, Kramvis A. Genotyping and molecular characterization of hepatitis B virus from human immunodeficiency virus-infected individuals in southern Africa. *PLoS One* 2012;7:e46345.
- [32] Baptista M, Kramvis A, Kew MC. High prevalence of 1762T 1764A mutations in the basic core promoter of hepatitis B virus isolated from black africans with hepatocellular carcinoma compared with asymptomatic carriers. *Hepatology* 1999;946–53. <https://doi.org/10.1002/hep.510290336>.
- [33] Andersson MI, Maponga TG, Ijaz S, Barnes J, Theron GB, Meredith SA, et al. The epidemiology of hepatitis B virus infection in HIV-infected and HIV-uninfected pregnant women in the Western Cape, South Africa. *Vaccine* 2013;5579–84.
- [34] Selabe SG, Lukhwareni A, Song E, Leeuw YGM, Burnett RJ, Mphahlele MJ. Mutations associated with lamivudine-resistance in therapy-naïve hepatitis B virus (HBV) infected patients with and without HIV co-infection: implications for anti-retroviral therapy in HBV and HIV co-infected South African patients. *J Med Virol* 2007;79:1650–4.
- [35] Mak D, Kramvis A. Molecular characterization of hepatitis B virus isolated from black South African cancer patients, with and without hepatocellular carcinoma. *Arch Virol* 2020;165:1815–25.
- [36] Jaquet A, Tchounga B, Tanon A, Bagny A, Ekouevi DK, Traore HA, et al. Etiology of hepatocellular carcinoma in West Africa, a case-control study. *Int J Cancer* 2018;143:869–77.
- [37] Kew MC. Synergistic interaction between aflatoxin B1 and hepatitis B virus in hepatocarcinogenesis. *Liver Int* 2003;23:405–9.
- [38] Cohen D, Ghosh S, Shimakawa Y, Ramou N, Garcia PS, Dubois A, et al. Hepatitis B virus preS2Δ38–55 variants: a newly identified risk factor for hepatocellular carcinoma. *JHEP Rep* 2020;2:100144.
- [39] Hwan YJ, Shin J, Jin Y-J, Lee J-W. Comparison of clinical manifestations and outcomes of noncirrhotic and cirrhotic hepatocellular carcinoma patients with chronic hepatitis B. *Eur J Gastroenterol Hepatol* 2020;32(1):66–73.
- [40] Cohen D, Ghosh S, Shimakawa Y, Ramou N, Garcia PS, Dubois A, et al. Hepatitis B virus preS2Δ38–55 variants: a newly identified risk factor for hepatocellular carcinoma. *JHEP Rep* 2020;2:100144.
- [41] Liu W-C, Wu I-C, Lee Y-C, Lin C-P, Cheng J-H, Lin Y-J, et al. Hepatocellular carcinoma-associated single-nucleotide variants and deletions identified by the use of genome-wide high-throughput analysis of hepatitis B virus. *J Pathol* 2017;243:176–92.
- [42] Suzuki Y, Maekawa S, Komatsu N, Sato M, Tatsumi A, Miura M, et al. HBV preS deletion mapping using deep sequencing demonstrates a unique association with viral markers. *PLoS One* 2019;14:e0212559.
- [43] Su I-J, Wang LH-C, Hsieh W-C, Wu H-C, Teng C-F, Tsai H-W, et al. The emerging role of hepatitis B virus Pre-S2 deletion mutant proteins in HBV tumorigenesis. *J Biomed Sci* 2014;21(1):98.
- [44] Pley C, Lourenço J, McNaughton AL, Matthews PC. Spacer domain in Hepatitis B virus polymerase: plugging a hole or performing a role? *J Virol* 2022;96:e0005122. <https://doi.org/10.1128/jvi.00051-22>.
- [45] Yang Z, Zhuang L, Lu Y, Xu Q, Tang B, Chen X. Naturally occurring basal core promoter A1762T/G1764A dual mutations increase the risk of HBV-related hepatocellular carcinoma: a meta-analysis. *Oncotarget* 2016;7:12525–36.
- [46] Zheng J-X, Zeng Z, Zheng YY, Yin S-J, Zhang D-Y, Yu Y-Y, et al. Role of hepatitis B virus base core and precore/core promoter mutations on hepatocellular carcinoma in untreated older genotype C Chinese patients. *J Viral Hepat* 2011;18:e423–31.
- [47] Skelton M, Kew MC, Kramvis A. Distinct mutant hepatitis B virus genomes, with alterations in all four open reading frames, in a single South African hepatocellular carcinoma patient. *Virus Res* 2012;163:59–65.
- [48] Maponga TG, McNaughton AL, van Schalkwyk M, Hugo S, Nwankwo C, Taljaard J, et al. Treatment advantage in HBV/HIV coinfection compared to HBV monoinfection in a South African cohort. *J Infect* 2020;81:121–30.
- [49] Wang M-L, Wu D-B, Tao Y-C, Chen L-L, Liu C-P, Chen E-Q, et al. The truncated mutant HBsAg expression increases the tumorigenesis of hepatitis B virus by regulating TGF-β/smād signaling pathway. *Virol J* 2018;15:61.
- [50] Qu L-S, Liu J-X, Liu T-T, Shen X-Z, Chen T-Y, Ni Z-P, et al. Association of hepatitis B virus pre-S deletions with the development of hepatocellular carcinoma in Qidong, China. *PLoS One* 2014;9:e98257.