


Identification of hepatitis B virus genotype I in Thailand

Vera Holzmayr¹ | Robert Hance² | Patricia Defechereux² | Robert Grant² |
Mary C. Kuhns¹ | Gavin Cloherty¹ | Mary A. Rodgers¹ 

¹Abbott Laboratories, Infectious Disease Research, Abbott Park, Illinois

²University of California San Francisco, San Francisco, California

Correspondence

Mary A. Rodgers, Abbott Laboratories, Infectious Disease Research, 100 Abbott Park Rd, Abbott Park, IL 60064.
Email: mary.rodgers@abbott.com

Funding information

National Institute of Allergy and Infectious Diseases, Grant/Award Numbers: UL1 RR024131, UO1-AI64002, RO1-AI118575, RO1-AI062333, T32-AI007641; Gladstone Institutes; Abbott Laboratories; Bill and Melinda Gates Foundation

The rare hepatitis B virus genotype I (HBV-I) classification includes complex A/G/C/U recombinants identified amongst the individuals from China, India, Laos, and Vietnam. Herein we report the first HBV-I specimen from Thailand, with detectable HBsAg despite a 10-amino-acid truncation. This HBV-I genome has a similar recombinant pattern to reference strains, including a C region that branches basal to references, suggesting a premodern era recombination event gave rise to HBV-I. With an average sequence divergence from other genotypes ranging from 7.66% (standard deviation [SD], 0.42%; C) to 14.27% (SD, 0.31%; H), this new genome supports the HBV-I classification as a unique genotype.

KEYWORDS

hepatitis B virus, recombinant virus, recombination, virus classification

1 | INTRODUCTION

Despite introduction of a vaccine for hepatitis B virus (HBV) in 1982, the global burden of HBV remains high, with an estimated 257 million chronic infections globally, of which more than 90% remained undiagnosed in 2015.¹ Accurate diagnosis of HBV is a critical initial step in caring for patients with HBV and preventing new transmissions; however, the diversity of HBV sequences presents a challenge for diagnostic tests.² It has been estimated that HBV was first introduced into the human population approximately 33 600 years ago,³ and since then it has evolved to nine unique genotypes (A-I) and a growing list of more than 30 subgenotypes.⁴⁻⁶ This diversity is largely driven by the error-prone viral reverse transcription polymerase, which lacks proofreading abilities and intergenotypic recombination.^{6,7} HBV genotypes and subgenotypes are unevenly distributed, with each HBV genotype predominating in unique regions of the world.⁴ In particular, HBV genotype I (HBV-I) has been found in China,⁸⁻¹¹ India,^{12,13} Laos,¹⁴ and Vietnam,¹⁵⁻¹⁷ with 39 unique sequences confirmed cases in Asia to date.⁵ In addition to these cases, HBV-I strains have also been sequenced from Vietnamese immigrants living in Canada¹⁸ and France,¹⁹ indicating that a more widespread distribution may exist for HBV-I outside of the Asia. Although HBV-I sequences branch together in a phylogenetic tree, recombination analysis has determined that HBV-I sequences are complex recombinants of genotypes A, C, G, and

unclassifiable sequences with exact breakpoints varying between individual strains.^{7,8,12,14-16,18,20} Classification of these strains as HBV-I or A/C/G/U recombinants is a topic of debate^{7,8,21,22} since the classical definition of a unique HBV genotype is at least 8% intergenotypic divergence, yet HBV-I strains are differ from the closest relative, genotype C, by only 7 to 7.9%.^{8,12,14,15,18,22} Thus, in additional, putative HBV-I genome sequences will be required to determine the appropriate classification for these unique sequences.

HBV coinfection with human immunodeficiency virus (HIV) puts patients at higher risk for severe liver disease, making prevention of such infections a priority.¹ The recent evaluation of HIV pre-exposure prophylaxis (PrEP) regimens by several groups has led to the successful implementation of PrEP programs amongst high-risk populations.^{23,24} Emerging evidence supports the safe use of PrEP amongst HBV carriers.²⁵ In a recent assessment of the safety of PrEP for patients with active HBV infection, serial blood samples were collected from 12 participants who had chronic HBV infections at the time of enrollment to monitor their infections with and without PrEP treatment during the iPrEx study.²⁵ HBV sequences obtained from these individuals indicated that resistance mutations did not develop amongst the participants on PrEP, supporting the safe use of PrEP for chronic HBV patients.²⁵ The samples were collected from diverse geographical regions (United States, Thailand, Peru, Ecuador, and South Africa), although the HBV genotypes for these infections were unknown. To evaluate the HBV diversity encompassed within the iPrEx study, we characterized the sequences

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. *Journal of Medical Virology* Published by Wiley Periodicals, Inc.

generated from serial blood draws of the chronic HBV participants and identified an HBV-I infection among them, which we characterized in depth to expand the known diversity of the genotype I classification.

2 | MATERIALS AND METHODS

2.1 | Samples

The sample cohort has been described previously,²⁵ and a complete diagnostic profile of the HBV-I infection time course can be found in Figure 1C of Solomon et al.²⁵ Sequences were generated from leftover cryopreserved plasma collected at the screening visit (day 12), 0, 79, 112, 251, and 336 days timepoints. All specimens were collected with informed consent from donors as approved by the University of California San Francisco IRB.

2.2 | Sequencing

Subgenomic sequences were generated for the surface antigen/reverse transcriptase (SRT) region as previously described.²⁶ The HBV-I genome was amplified in two overlapping fragments using Advantage-2 polymerase reagents (Clontech, Palo Alto, CA). Primer pairs for the S-core fragment were HBV-238F (5'-ATACCACAGAGTCTAGACTCGTGGTGGACT-3', nucleotides (nt) 236-265)/ Hbc-2382R (5'-CGTCKGCGAGGCGAGGGAGTTC-3', nt 2382-2403) for the first-round PCR and HBrt-295F (5'-GTGTCYTGCCWAAATTCGCAGTCCC-3', nt 295-320)/Hbc-2298R (5'-TGTKGATARGATAGGGCA TTTGGTGGTC-3', nt 2298-2336) for the nested PCR. Primers for the core-pol fragment were Hbc-1776F (5'- GGAGGCTGTAGGCAT AAATTGGTCTG-3', nt 1776-1801)/ HBrt-1253R (5'- GCAGTAT GGATCGGCAGAGGAG-3', nt 1253-1274) for the first-round PCR and Hbc-1882F (5'- CCTTGGGTGGCTTTTRGRC-3', nt 1882-1901)/

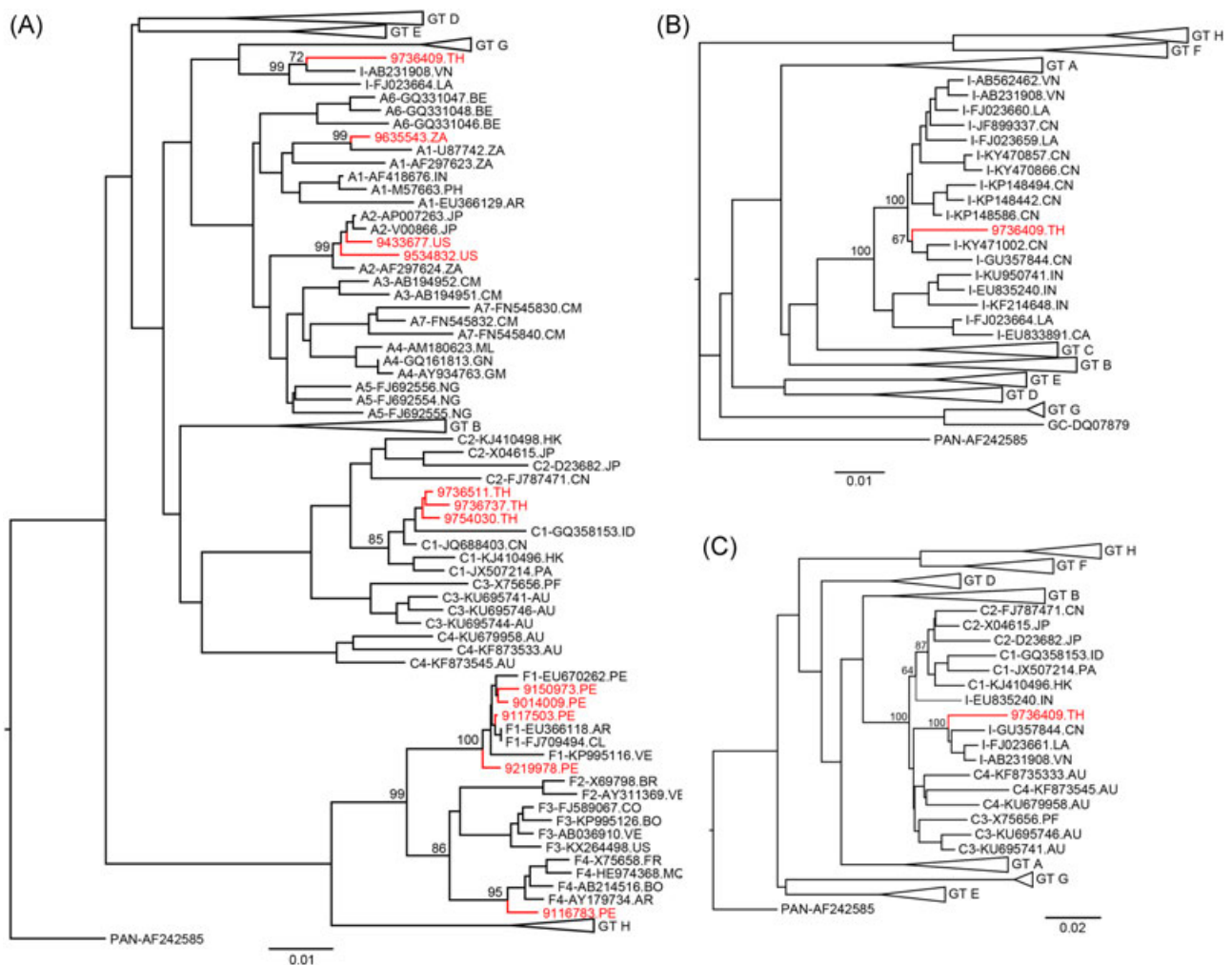


FIGURE 1 HBV phylogenetic trees. Neighbor-joining phylogenetic trees are shown for the 798 nucleotide surface antigen/reverse transcriptase (SRT) fragment (genome coordinates 204-1001 nucleotides) (A), the 3215 nucleotide complete genomes (B), and the nucleotide position 1500 to 2900 region of the HBV genome (C). Branches without any sample sequence are depicted as triangles. Sample sequences are indicated in red and references are in black. Individual strains are labeled with their classification, GenBank accession number, and the two-letter country code for their source. Bootstrap values are provided for nodes relevant for sample classification. The tree in (B), is representative of a larger tree that includes a comprehensive set of 37 HBV-I reference sequences from Canada, Vietnam, China, Laos, and India. Two-letter codes at the end of each sequence label indicate the source country. GT, genotype; HBV, hepatitis B virus; HBV-I, hepatitis B virus genotype I

HBrt-1186R (5'-CCAGTGGGGTTCGRTCAGC-3', nt 1186-1205) for nested PCR. First-round and nested second-round amplifications consisted of preincubation at 95°C for 1 minute, 40 cycles of 95°C for 15 seconds and combined annealing/extension at 68°C for 5 minutes, and final extension at 68°C for 10 minutes. Sanger sequence data was analyzed and assembled using Sequencher v 5.4.1 (Gene Codes Corp, Ann Arbor, MI). HBV surface antigen (HBsAg) subtyping was completed as previously described²⁶ and the Geno2Pheno HBV online tool (Max Planck Institut Informatik, Germany; <http://hbv.geno2pheno.org>) was used to identify escape and resistance mutations.

2.3 | Phylogenetic analysis

Sequences were merged with reference sequence alignments including 29 complete HBV genotype I genomes obtained from a comprehensive alignment generated from the Genbank database on 13 February 2018 (version 1.0 alignment from Bell et al⁵), plus eight additional published complete genome sequences^{14,15,18} in BioEdit v7.2.5.²⁷ Neighbor-joining HBV phylogenetic trees were generated as previously described,²⁶ and trees were visualized using FigTree software (version 1.4.2; A. Rambaut, Institute of Evolutionary Biology, University of Edinburgh, Scotland). HBV-I strains branching closely with other references were removed to generate a simpler tree for visualization in Figure 1B. Recombinant analysis was completed using Simplot (version 3.5.1) using a window of 400 basepairs (bp), step size of 20 bp, GapStrip on, 100 reps, Kimura two-parameter, T/t 2.0, and neighbor-joining settings. Genetic distances were calculated in BioEdit with complete genotype A to H reference strain sequence alignments generated from the Genbank database on 26 June 2017 (version 0.4 alignment from Bell et al⁵).

2.4 | Serology testing

The HBV-I specimen collected at the study entry day 0 was tested at a 1:500 dilution using ARCHITECT HBsAg Qualitative II, and the ARCHITECT CORE assays, and neat using the ARCHITECT HBeAg assay in combination with the ARCHITECT HBeAg Quantitative Calibrators (Abbott Diagnostics, Abbott Park, IL.)

2.5 | Data availability

The HBV-I genome has been deposited into GenBank under accession MH368022. Subgenomic HBV sequences for genotypes A, C, and F specimens have been deposited under accessions MH368023-MH368033.

3 | RESULTS

Phylogenetic classification of SRT sequences obtained from plasma donations of 12 chronic HBV carriers enrolled in the iPrEx study identified genotype A1 (South Africa), A2 (United States), C1 (Thailand), F1 (Peru), F4 (Peru), and I (Thailand) infections

(Figure 1A). Given the limited number of genome sequences available for HBV-I infections, two additional overlapping amplicons were sequenced from the HBV-I specimen to produce a complete genome of 3215 nucleotides. This specimen was collected in Chiang Mai, Thailand, 12 days before patient 9736409 began PrEP treatment, and subgenomic sequences generated from serial samples collected at later timepoints during PrEP treatment were all identical to the sequence obtained from the initial sample. Phylogenetic analysis of the complete genome indicated that the 9736409 sequence branched with other HBV-I reference strains, but was basal to a set of reference strains from China (Figure 1B). SimPlot confirmed that the 9736409 sequence shared 95.1% to 99% similarity to a genotype I consensus sequence across the entire length of the genome (Figure 2A), and recombinant bootscan analysis indicated that it had a U/C recombinant pattern similar to other HBV-I genomes from China and Vietnam (Figure 2B-D),^{8,12,14,15,18,22} with a genotype C region at nucleotide positions 1500 to 2900. Within the HBV surface antigen open reading frame (HBsAg ORF), the HBsAg subtype was determined to be *adw2* and an early termination mutation was identified at amino acid 216 (T801A). Likewise, an early stop codon was identified in the precore region of the 9736409 sequence at amino acid 28 (G1896A), which encodes the HBV e antigen (HBeAg). Importantly, no resistance mutations were found in the complete RT ORF as previously noted.²⁵ The average nucleotide divergence of the 9736409 genome ranged from 7.66% (standard deviation [SD], 0.42%) for $n = 1543$ genotype C reference sequences to 14.27% (SD, 0.31%) for $n = 17$ genotype H reference sequences. In contrast, 9736409 was only 3.06% (SD, 0.52%) divergent compared to $n = 29$ unique genotype I strains with complete genome sequences. The diluted HBV-I specimen was HBsAg positive by ARCHITECT HBsAg Qualitative II (3269.79 S/CO), and anti-HBc positive (8.25 S/CO). The same specimen was HBeAg negative when tested neat.

4 | DISCUSSION

The unusual complex recombinants classified as HBV-I include a growing number of sequences from individuals born in Asia, where HBV is highly endemic.⁴ The HBV-I genotype has been defined by strains having 7% to 8% nucleotide divergence to the closest relative genotype C^{8,12,14,15,18,22} and the new recombinant genome sequence presented here, 9736409, meets this criteria with a mean nucleotide divergence of 7.66% (SD, 0.42%) to genotype C reference sequences. Therefore, this sequence is classified as genotype I, making it the first HBV-I strain identified from Thailand. With a HBsAg serotype of *adw2*, this specimen meets the criteria for the I1 subgenotype.²⁸ Like other sequences classified as genotype I, 9736409 is a complex U/C recombinant with a similar recombination pattern and breakpoints (1500-2900) as other strains from China and Vietnam (Figure 2).^{8,12,14,15,18,22} Although other HBV-I strains also have A- and G-like regions¹³⁻¹⁷ (Vietnam refs, Laos, India), these segments could not be classified in the 9736409 genome due to low bootstrap support (Figure 2). Notably, genotype G sequences are quite rare in

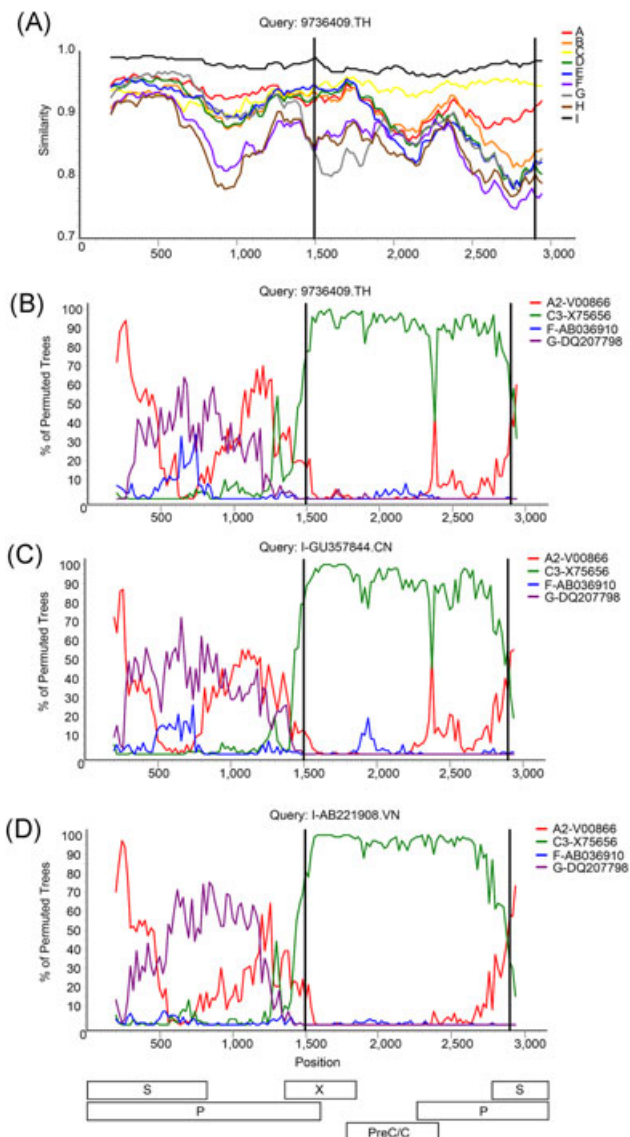


FIGURE 2 Recombinant analysis. A similarity plot generated in SimPlot with consensus sequences for each genotype is shown in panel (A). Bootscan plots comparing sample 9736409 (B), or HBV-I strains from China (C), and Vietnam (D), to indicated reference strains are shown above a diagram of the HBV ORFs corresponding to the nucleotide positions indicated by the x-axis. Breakpoints are indicated by a black line. HBV, hepatitis B virus; HBV-I, hepatitis B virus genotype I; ORFs, open reading frames

Asia, with all reported strains coming primarily from the United States, France, and Germany.⁴ However, a genotype G/C recombinant (CU400) has been reported in Thailand,²⁹ indicating that genotype G strains might be present in the region. Yet the 9736409 strain is unlikely to be closely related to the CU400 strain since the C-regions of the two sequences have different breakpoints and the two strains only share 89% sequence identity.²⁹ Consistent with previous HBV-I genome analyses,^{8,13,16} the region of the 9736409 genome classified as genotype C branched basal to reference strains (Figure 1C). This suggests that the recombination events that resulted in HBV-I did not occur recently, although in

addition, analysis and sequences will be required to explore the origins of HBV-I further. Likewise, the unclassifiable region from 1 to 1500 (Figure 2) is distinct from all other HBV genotypes (Figure 2),⁸ which further supports the identification of I as a separate genotype.

Examination of the amino acid sequences of the 9736409 open reading frames (ORFs) did not identify escape or resistance mutations in HBsAg or the polymerase. However, an early stop codon was identified in the HBsAg ORF at amino acid position 216 (T801A), which is a naturally occurring variant that results in a 10-amino-acid truncation of the HBsAg ORF.³⁰ While our direct amplicon sequencing method identified this variant as the major sequence present in 9736409, the possibility cannot be ruled out that a minor population of wildtype virus could produce full length HBsAg in this specimen. In a comparison of all 39 HBV-I genomes with HBsAg and pre-C coverage, only one other strain (LA.KF214679, G668A) carried a nonsense HBsAg mutation, although it is unknown whether HBsAg was detectable in this specimen.¹³ Similarly, a minority carried pre-C nonsense mutations ($n = 7$; LA.KF214679, LA.KF214648, LA.KF214680, LA.FJ023665, LA.FJ023666, LA.FJ023672, and IN.EU835242), indicating that the nonsense mutations identified in the 9736409 genome are not representative of the entire genotype. Nonetheless, despite an early HBsAg stop codon and unclassifiable sequence in the HBsAg ORF, an HBsAg level of 34 456 IU/mL was detected by the ARCHITECT HBsAg assay (Abbott Diagnostics) and a viral load of 3.26 log IU/mL was detected by the m2000 RealTime HBV assay (Abbott Molecular Diagnostics, Des Plaines, IL) in the 9736409 specimen collected upon enrollment in the iPrEx study.²⁵ Furthermore, the specimen was HBeAg negative as expected due to the precore amino acid 28 stop codon (G1896A). As HBV sequence diversity continues to expand, persistent global HBV surveillance remains essential to ensuring diagnostic tests can detect all strains regardless of their sequence or geographical location.

ACKNOWLEDGMENTS

The authors are grateful to the entire iPrEx study team and participants for their contributions to this study. In particular, the authors acknowledge Pongpun Saokhieo and Dr Suwat Chariyalertsak from the iPrEx team in Thailand for collecting the HBV-I specimen. The authors also thank Dr Catherine Brennan for initiating this project and Barbara Harris for technical input.

ORCID

Mary A. Rodgers  <http://orcid.org/0000-0001-8815-8651>

REFERENCES

- World Health Organization. *Global hepatitis report*. Geneva: WHO; 2017.
- Alavian SM, Carman WF, Jazayeri SM. HBsAg variants: diagnostic-escape and diagnostic dilemma. *J Clin Virol*. 2013;57(3):201-208.
- Paraskevis D, Magiorkinis G, Magiorkinis E, et al. Dating the origin and dispersal of hepatitis B virus infection in humans and primates. *Hepatology*. 2013;57(3):908-916.

4. Sunbul M. Hepatitis B virus genotypes: global distribution and clinical importance. *World J Gastroenterol.* 2014;20(18):5427-5434.
5. Bell TG, Yousif M, Kramvis A. Bioinformatic curation and alignment of genotyped hepatitis B virus (HBV) sequence data from the GenBank public database. *SpringerPlus.* 2016;5(1):1896.
6. Kay A, Zoulim F. Hepatitis B virus genetic variability and evolution. *Virus Res.* 2007;127(2):164-176.
7. Araujo NM. Hepatitis B virus intergenotypic recombinants worldwide: an overview. *Infect Genet Evol.* 2015;36:500-510.
8. Fang ZL, Hue S, Sabin CA, et al. A complex hepatitis B virus (X/C) recombinant is common in Long An county, Guangxi and may have originated in southern China. *J Gen Virol.* 2011;92(Pt 2):402-411.
9. Yu H, Yuan Q, Ge SX, et al. Molecular and phylogenetic analyses suggest an additional hepatitis B virus genotype "I". *PLOS One.* 2010;5(2):e9297.
10. Tong W, He J, Sun L, He S, Qi Q. Hepatitis B virus with a proposed genotype I was found in Sichuan province, China. *J Med Virol.* 2012;84(6):866-870.
11. Zhou TC, Li X, Li L, Li XF, Zhang L, Wei J. Evolution of full-length genomes of HBV quasispecies in sera of patients with a coexistence of HBsAg and anti-HBs antibodies. *Sci Rep.* 2017;7(1):661.
12. Arankalle VA, Gandhe SS, Borkakoty BJ, Walimbe AM, Biswas D, Mahanta J. A novel HBV recombinant (genotype I) similar to Vietnam/Laos in a primitive tribe in eastern India. *J Viral Hepat.* 2010;17(7):501-510.
13. Haldipur BP, Walimbe AM, Arankalle VA. Circulation of genotype-I hepatitis B virus in the primitive tribes of Arunachal Pradesh in early sixties and molecular evolution of genotype-I. *Infect Genet Evol.* 2014;27:366-374.
14. Olinger CM, Jutavijittum P, Hübschen JM, et al. Possible new hepatitis B virus genotype, southeast Asia. *Emerg Infect Dis.* 2008;14(11):1777-1780.
15. Huy TTT, Ngoc TT, Abe K. New complex recombinant genotype of hepatitis B virus identified in Vietnam. *J Virol.* 2008;82(11):5657-5663.
16. Norder H, Hannoun C, Lindh M. An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. *J Gen Virol.* 2000;81(Pt 9):2267-2272.
17. Nguyen CH, Ishizaki A, Chung PTT, et al. Prevalence of HBV infection among different HIV-risk groups in Hai Phong, Vietnam. *J Med Virol.* 2011;83(3):399-404.
18. Osiowy C, Kaita K, Solar K, Mendoza K. Molecular characterization of hepatitis B virus and a 9-year clinical profile in a patient infected with genotype I. *J Med Virol.* 2010;82(6):942-948.
19. Colson P, Roquelaure B, Tamalet C. Detection of a newly identified hepatitis B virus genotype in southeastern France. *J Clin Virol.* 2009;45(2):165-167.
20. Ismail AM, Goel A, Kannangai R, Abraham P. Further evidence of hepatitis B virus genotype I circulation in Northeast India. *Infect Genet Evol.* 2013;18:60-65.
21. Kurbanov F, Tanaka Y, Mizokami M. Geographical and genetic diversity of the human hepatitis B virus. *Hepatol Res.* 2010;40(1):14-30.
22. Pourkarim MR, Amini-Bavil-Olyae S, Kurbanov F, Van Ranst M, Tacke F. Molecular identification of hepatitis B virus genotypes/subgenotypes: revised classification hurdles and updated resolutions. *World J Gastroenterol.* 2014;20(23):7152-7168.
23. Amico KR, Mehrotra M, Avelino-Silva VI, et al. Self-reported Recent PrEP dosing and drug detection in an open label PrEP study. *AIDS Behav.* 2016;20(7):1535-1540.
24. Riddell J, Amico KR, Mayer KH. HIV preexposure prophylaxis: a review. *JAMA.* 2018;319(12):1261-1268.
25. Solomon MM, Schechter M, Liu AY, et al. The safety of tenofovir-emtricitabine for HIV pre-exposure prophylaxis (PrEP) in individuals with active hepatitis B. *J Acquir Immune Defic Syndr.* 2016;71(3):281-286.
26. Harris BJ, Holzmayr V, Qureshi H, et al. Hepatitis B genotypes and surface antigen mutants present in Pakistani blood donors. *PLOS One.* 2017;12(6):e0178988.
27. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 1999;41:95-98.
28. Kramvis A. Genotypes and genetic variability of hepatitis B virus. *Intervirology.* 2014;57(3-4):141-150.
29. Suwannakarn K, Tangkijvanich P, Theamboonlers A, Abe K, Poovorawan Y. A novel recombinant of Hepatitis B virus genotypes G and C isolated from a Thai patient with hepatocellular carcinoma. *J Gen Virol.* 2005;86(Pt 11):3027-3030.
30. Chaouch H, Taffon S, Villano U, et al. Naturally occurring surface antigen variants of hepatitis B virus in Tunisian patients. *Intervirology.* 2016;59(1):36-47.

How to cite this article: Holzmayr V, Hance R, Defechereux P, et al. Identification of hepatitis B virus genotype I in Thailand. *J Med Virol.* 2019;91:717-721. <https://doi.org/10.1002/jmv.25346>