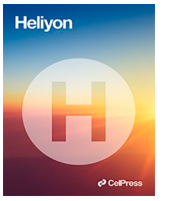
Heliyon 9 (2023) e22358

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Characterization of Hepatitis B virus based complete genome analysis improves molecular surveillance and enables identification of a recombinant C/D strain in the Netherlands

Jeroen Cremer, Francoise van Heiningen, Irene Veldhuijzen, Kimberley Benschop *

Center for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands

ARTICLE INFO

Keywords:
Hepatitis B virus
Phylogenetic analysis
Complete genome
S-gene
C-gene
Recombinants

ABSTRACT

Hepatitis B Virus (HBV) is classified into 10 HBV genotypes (A-J) based a >7.5 % divergence within the complete genome or a >4 % divergence in the S-gene. In addition, recombinant strains with common breakpoints at the gene boundaries of the preS1/preS2/S- and preC/C-gene are often identified. Analysis of HBV based on the complete genome is essential for public health surveillance as it provides higher genetic resolution to conduct accurate characterization and phylogenetic analysis of circulating strains and identify possible recombinants. Currently two separate assays are used for HBV-surveillance; the S-gene for typing, and due to the higher genetic variation, the C-gene to gain insight in transmission patterns.

The aim of the study was to develop a complete genome PCR-assay and evaluate the characterization and circulation of HBV strains through the use of the S-gene, C-gene and complete genome. For this HBV positive samples collected in the period 2017 through 2019 were selected.

Analysis of the complete genome showed that complete genome analysis portrays a high genetic resolution that provided accurate characterization and analysis of the different circulating types in the Netherlands and enabled identification and characterization of a recombinant CD strain.

1. Introduction

An estimated 250–260 million people in the world are reported to be Hepatitis B surface Antigen (HBsAg) positive [1] and it remains a global public health concern. The WHO has set course on the elimination of HBV in 2030 by active case finding, contact tracing and treatment and vaccination plans.

HBV is a double stranded DNA virus with a genome of ~3.2 kb in length. The genome contains four genes that codes for seven different proteins within overlapping open reading frames (Fig. 1). The S-gene is preceded by a preS1 and S2-gene, which result in the large, medium and small for of the surface antigen (HBsAg). The C-gene encodes the core antigen (HBcAg), while in combination with the preC-gene encodes the HBV e antigen (HBeAg). Finally the P-gene encodes the polymerase and the X-gene encodes the transcriptional transactivator HBV X protein (HBx).

The genetic variability of HBV is very high due to the lack of the proof-reading mechanism. HBV is classified in 10 HBV genotypes

* Corresponding author. Laboratory for Infectious Diseases and Screening, Center for Infectious Diseases National Institute of Public Health and the Environment, Pobox 1, 3720 BA, Bilthoven, the Netherlands.
E-mail address: kim.benschop@rivm.nl (K. Benschop).

https://doi.org/10.1016/j.heliyon.2023.e22358
Received 18 August 2023; Received in revised form 23 October 2023; Accepted 10 November 2023
Available online 14 November 2023
2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

designated A-J, based on the level of divergence of the complete genome (>7.5 % divergence in total or >4 % divergence in the S-gene) [3–6]. The different types have a distinct geographical distribution with genotype A commonly found in North America and Europe and genotype D in North Africa and the Middle East. The dominant types in the Netherlands have been reported to be genotype A and D [7,8]. The different types also differ in disease progression and clinical outcome as well as reactivity to antiviral therapy [6,9–11].

In addition to the genetic variability, HBV can undergo frequent recombination events adding to the variability of these viruses. More recent recombinants include combinations of genotypes A/B/C, A/C, A/D, A/E, A/D, A/G, C/G, B/D, C/F, C/G, C/J, A/C/G, D/F, F/G, D/F, A/F [12–14]. Common breakpoints are characterized at the gene boundaries of the preS1/preS2/S and preC/C-gene [13,15].

The worldwide distribution of various recombinants shows the ability of these viruses to spread with distinct distributions and epidemiology [13,15]. Other studies have indicated recombination events as an evolutionary step to enhance pathogenicity [16] and may therefore have potential clinical implications [12]. As recombination finds its basis in combining genetic material from different genotypes, typing of only one region of the HBV genome can lead to incorrect genotype assignment [3,12,17]. As such, analysis of complete genome of HBV is critical for identification and accurate characterization of recombinant strains.

The variation of HBV, may it be due to genetic evolution, acquirement of substitutions, insertion, deletions, or recombination, requires accurate characterization and analysis of circulating strains for public health surveillance and analysis of complete genome becomes a prerequisite for proper HBV surveillance.

While HBV classification is often based on the S-gene (>4 % divergence), the C-gene is used to gain insight in transmission patterns due to it higher divergence [2]. However, the C-gene does not distinguish between genotype B and C and genotypes D and E, due to recombination events in the past [15,18,19].

In this study, we developed a Sanger based PCR-sequencing assay to generate complete genomes of HBV and investigated the value of complete genome analysis for characterization and analysis of HBV strains versus analysis of the S-gene for typing and analysis of the C-gene for transmission analysis separately among HBV strains from Dutch HBV infected cases.

2. Materials and methods

2.1. Samples

In this study, HBV positive samples collected in the period 2017 through 2019 were selected. Samples were obtained from acute and chronic cases notified to OSIRIS by the Municipal Health Services (MHS). OSIRIS is an anonymized national electronic database used for infectious disease surveillance. For some cases acute or chronic status was not disclosed and were categorized as "unknown" status. All acute cases diagnosed in the Netherlands with a positive HBsAg test result and/or an anti-HBc IgM result are notifiable. Chronic cases are notifiable and reported when a positive HBsAg or HBV-DNA positivity is diagnosed in the Netherlands for the first time. Interviews are conducted with these cases by the MHS to ascertain risk exposures and enable source-tracing investigations and/or partner notification. Samples are requested for typing from all acute cases, and from chronic cases with risk behavior (i.e. tested as part of the HBV vaccination program for behavioral risk groups, or with reported transmission route MSM sexual contact). Available samples are sent to the RIVM for genotyping.

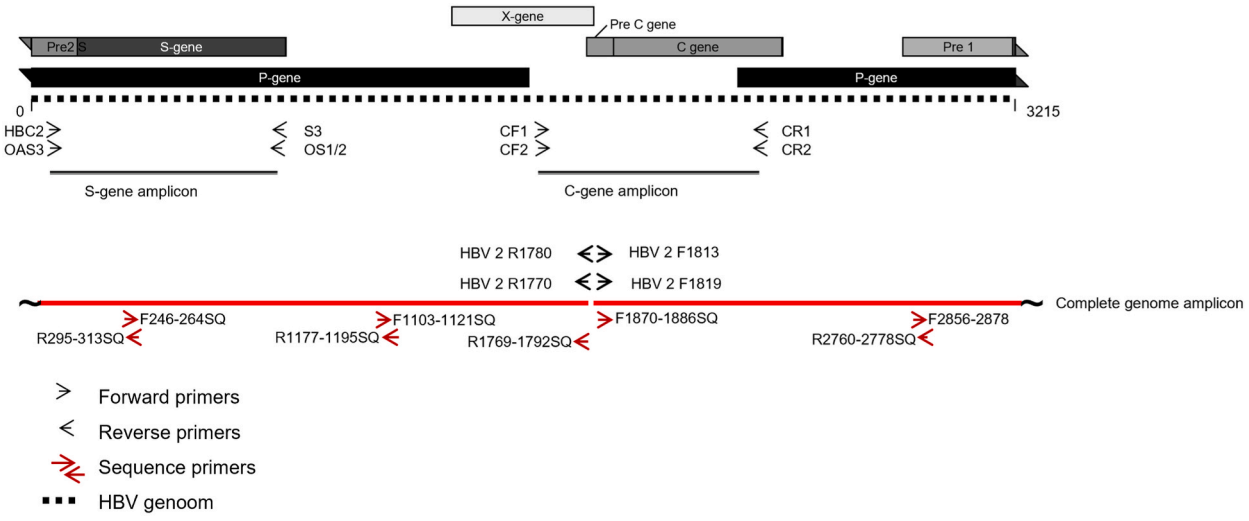


Fig. 1. Schematic representation of the complete genome assay, alongside the previously published S- and C-gene assay. Position of primers is based on reference genome genotype A (accession number X02763). Adapted from Ref. [2].

2.2. Characterization of complete genome of HBV

The complete genome was amplified in a nested PCR approach (Table 1, Fig. 1) generating a first amplicon of 3216 bp and a second amplicon of 3200 bp. DNA was isolated from 200 μL of serum by automated extraction by using the LC Nucleic Acid isolation kit (Roche, Almere, the Netherlands). Isolated DNA was eluted in 50 μL of elution buffer. The first reaction (25 μl) contained 2.5 μl of isolated DNA, 1× reaction buffer 1.5 mM of MgCl2, 0.4 pmol of the forward and reverse primer each (HBV1–F1813 and HBV2-R1780, Table 1), 0.3 mM dNTP's, and 1 U of Silverstar polymerase (Eurogentic, Seraing). Amplification was carried out by heating the mixture 30 s for 95 °C, followed by 34 cycles of 95° for 16 s, 60 °C for 16 s and 70 °C for 90 s, and 1 cycle 70 °C for 160 s. The reaction was ended with a 4 °C extended hold. The second/nested reaction (50 μl) contained 3 μl of first PCR mix, 1× reaction buffer, 1.5 mM of MgCl2, 0.4 pmol of the forward and reverse primer each (HBV2–F1770 and HBV1–F1819, Table 1), 0.3 mM dNTP's, and 2 U of Silverstar polymerase (Eurogentic, Seraing). Amplification was carried out by heating the mixture 32 s for 95 °C, followed by 36 cycles of 95° for 16 s, 60 °C for 16 s and 70 °C for 90 s, and 1 cycle 70 °C for 160 s. The reaction was ended with a 4 °C extended hold. Amplicons were purified using EXosapit (Ampliqon; Denmark) and the sequence reaction was done by BaseClear (Leiden) with Big Dye Terminator (Life Technologies) using HBV2–F1770 and HBV1–F1819 primers (Table 1). The lower limit of detection was determined using the Hepatitis B standard Nap-HBV006 (Acromertix, Genotype A1, 2e7 IU/ml) in triplicate.

2.3. Molecular analysis

The generated sequences were analyzed and assembled with BioNumerics (7.6.3) (Sint-Martens-Latem, Belgium). The generated HBV sequences were aligned with Simmonic Sequence Editor (SSE) [20]. As the sequence ends are not always successful, and to assure that all sequences comprise the same region, we analyzed sequence data from nucleotide 1916 (C-gene) to 1751 (X-gene) (nucleotide numbering based on gH reference genome, AB275308). Genotype were assigned based on the similarity analysis of the complete genome with a reference set of HBV types (A-J) assigned by Pourkarim M.R. et al., 2014 [3] using Bionumerics version 7.1. Genotype was assigned based on S-gene divergence of 4 % and complete genome divergence of 7.5 % to a reference genotype data set [21]. Phylogenetic trees were constructed with the Nextstrain platform [22](https://nextstrain.org/)). In short, sequences were aligned by using MAFFT [23], a phylogenetic tree was inferred with IQ-TREE [24], and time-resolved trees were generated with TreeTime [25], which estimates the mutation rate. In order to confirm the evolutionary divergence and clustering shown in Nextstrain, neighbour-joining trees with the model Maximum Composite Likelihood and 1000 bootstraps were constructed using MEGA 7.0 software [26]. Recombinant strain 17-126 was excluded from phylogenetic analysis as proposed by Pourkarim et al. [17]. Strains with a complete genome that were >95 % complete and between nucleotide positions 1916 (C-gene) to 1751 (X-gene) (numbering based on AB275308) were analyzed.

Mean nucleotide divergence of each genotype identified in this study was calculated with MEGA7.0 software [26] over the S-gene, C-gene and complete genome (within group divergence) using the p-distance model for nucleotide sequences. Statistical significance was calculated with the z-test (two tailed) and $P < 0.05$ was defined as significant.

Recombination was analyzed with Simplot and Bootscan software version 3.5.1.

2.4. GenBank

The near complete sequences were deposited in GenBank accession numbers: MW999475-MW999678.

Table 1
PCR Primers for sequencing of the near complete genome of HBV.

Primer Name	Sequence'5′-3	Position[a]
First PCR (Single FL PCR)		
HBV1–F1813	CCAKCACCATGCAACTTTTTCACCTC	1812–1834
HBV2-R1780	GMACAGACCMATTTATGCCTACAGC	1780–1805
Second PCR (Nested FL PCR)		
HBV2-R1770	ATGCCTACAGCCTCCTARTAC	1770–1796
HBV1–F1819	GCTTTTTCACCTCTGCCTARTCATCT	1816–1844
Sequencing		
F1870-1886SQ	CTCCAAGCTGTGCCTTG	1870–1886
F2856-2878	GGGTCACCHTATTCBTGGGAAC	2820–2841
F246-264SQ	GTCTAGACTCGTGGTGGAC	246–264
F1103-1121SQ	CAACTTACAAGGCCTTTCT	1102–1121
R2760-2778SQ	TCTGGAAGTAATGATTAAC	2724–2742
R295-313SQ	CGAATTTTGGCCAAGACAC	295–313
R1177-1195SQ	TTGCGTCAGCAAACACTTG	1177–1195
R1769-1792SQ	CCTACAGCCTCCTARTACA	1769–1787

[a] Position of primers is based on reference genome genotype A (accession number X02763).

Table 2
Demographic data and nucleotide divergence of Dutch HBV strains.

Genotype	N strains	Infection status			Median age in years (range)	% males	% nucleotide divergence $						p-value*		
		Acute	Chronic	Unknown			S-gene		C-gene		GC		S/C	S/CG	C/CG
							Mean	SD	Mean	SD	Mean	SD			
A	123	94	22	7	51 (18–78)	74.8	0.61	0.14	1.73	0.26	1.86	0.11	0.00	0.00	0.00
B	15	10	1	4	42 (32–77)	60.0	1.51	0.23	3.15	0.39	3.84	0.20	0.00	0.00	0.00
C	13	11	1	1	36 (20–67)	100	1.13	0.22	2.26	0.31	2.77	0.18	0.03	0.00	0.05
CD	1	1	0	0	57 (57–57)	0.0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
D	26	17	4	5	38 (22–68)	50.0	1.03	0.14	3.40	0.39	2.28	0.13	0.00	0.00	0.00
E	4	1	2	1	28 (28–72)	75.0	0.66	0.22	3.97	0.59	2.40	0.20	0.00	0.00	0.01
F	15	14	1	0	28 (20–53)	66.7	0.06	0.03	0.10*	0.05	0.13	0.03	0.03	0.00	0.09
H	1	1	0	0	29	100	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Total	198	149	31	18	47 (18–78)	71 %	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Genotype were assigned based on the similarity analysis of the complete genome with a reference set of HBV types (A-J) and subtypes published by Pourkarim MR et al., 2014 [3] using Bionumerics version 7.1.

$p-distance model for nucleotide sequences (MEGA7.0) of S-gene, C- gene, and complete genome (CG); SD = standard deviation.

* Statistical difference (Z-test with critical z-score values of 1.96; $P < 0.05$) between S-gene and C- gene, S-gene to complete genome and C-gene to complete genome.

na = not applicable.

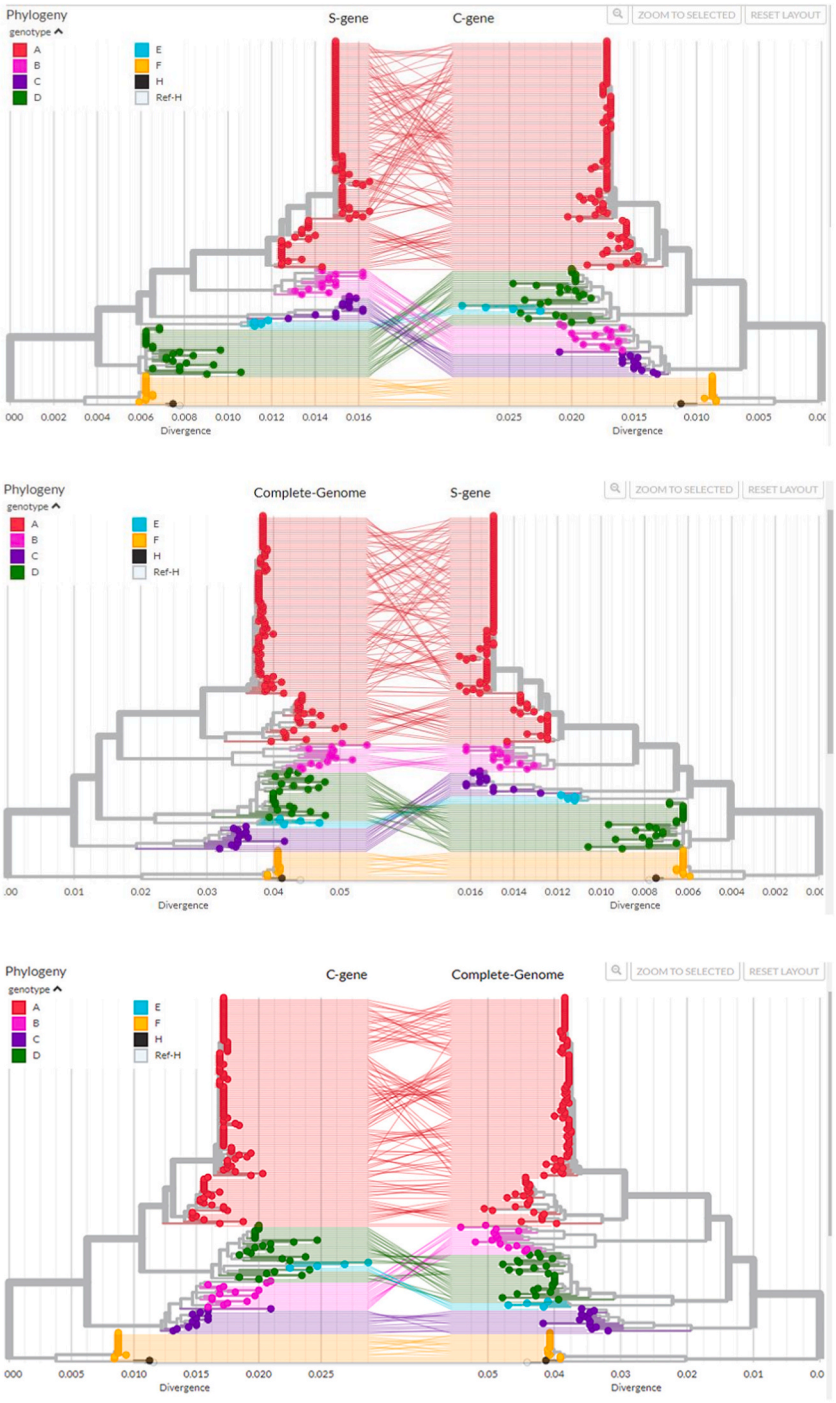


Fig. 2. Printscreen of the phylogenetic tanglegram analysis based on S-gene and C-gene (A), S gene and near complete genome (B), and C-gene and S-gene (C). Genotype was assigned based on the similarity analysis of the complete genome with a reference set of HBV types (A–J) published by Pourkarim MR et al., 2014 [3] using Bionumerics version 7.1 The tanglegram was generated by nextstrain with the study sequences (n = 198) labeled by genotype (A-red; B-blue; C-pink; D-purple; E-dark green; F-yellow; H-dark grey). The reference strain (AB275308-H) is indicated in light grey.

3. Results

3.1. Assay performance

The lower limit of detection for the complete genome assay was 500 IU/ml. A total of 263 HBsAg positive samples in the study period were analyzed using the complete genome assay. A total of 198 (79.8 %) HBV samples were successfully sequenced within the complete genome and were >95 complete. For 12 samples only a partial genome was sequenced, and for 5 samples the complete genome sequence was of poor quality. Complete genome sequencing was not achieved in 48 samples, of which 24 strains showed a ct value > 33. Of the other 24 sequences (ct 21.3–36.5) the partial S gene could be sequenced using the assay by Boot et al. [2].

3.2. Genotype and demographic data

Table 2 shows the infected genotypes and demographic data of the cases successfully sequenced with >95 complete genome between 2017 and 2019. On the basis of the complete genome, we identified genotype (g)A through F and H, and a recombinant genotype CD. The strains were isolated from cases with a reported acute infection (n = 149), cases from behavioral risk groups with a reported chronic infection (n = 31), and 18 cases where infection status was unknown. The overall median age was 47 years (range 18–78 years). Genotypes E and F, had the lowest median age of 28 years old (ranges 28–72 years and 20–53 years respectively), whereas genotype A showed the highest median age of 51 years (range 18–78 years).

A majority of cases were male for all genotypes, with the exception of genotype D, with 50 % of cases being male,

For 147 cases the country of birth was known, and of these, the Netherlands (n = 113 (76.9 %) was frequently reported as country of birth. For all other cases with country of birth reported, country of birth varied from various parts of the world with 1–2 cases (different genotypes) per country and 3 cases with genotype B from China.

3.3. Phylogenetic and divergence of S-gene versus C-gene versus complete genome

Fig. 2 shows the genetic resolution of the divergence based on the S-gene, the C-gene and complete genome as a tanglegram. Type specific clustering with a bootstrap >70%was confirmed by MEGA7. A higher resolution is found in the phylogenetic tree of the C-gene and complete genome compared to the S-gene due to the higher nucleotide divergence in this gene (Fig. 2, Table 2).

Mean within group divergence analysis shows the lowest genetic divergence to be within the S-gene for all genotypes compared to the C-gene and complete genome (Table 2).

For the S-gene, C-gene and complete genome, the mean within group nucleotide divergence among the Dutch strains is lowest for genotype F (0.06 %, 0.10 %, and 0.13 %, respectively, n = 15 strains) (Table 2), followed by genotype A (0.61 %, 1.73 %, and 1.86 %, respectively, n = 123 strains). Mean within group nucleotide divergence based on the S-gene and the complete genome is highest for genotype B (1.51 % and 3.84 % respectively, n = 15 strains). However, based on the C-gene, the mean within group nucleotide divergence is highest for genotype E (3.97 %, n = 4 strains) and genotype D (3.40 %, n = 26 strains).

The higher divergence in the C-gene and complete genome compared to the S-gene was statistically significant for all genotypes. Of interest, there was a significant difference also found for genotypes A through E between the C-gene and complete genome, while for genotype F this was not significant (Table 2).

Within the C-gene phylogenetic reconstruction and typing for genotype D-E and B–C are hampered due to intragenotypic recombination between these strains in the past (Figs. 2A and 1C) [15]. Typing of genotype D and E is indeed hampered within the C-gene, clustering together with a bootstrap of 93 %, but can be easily distinguished when analyzing the complete genome (Figs. 2B and 1C). In contrast, in our dataset the genotype B and C can be distinguished based on divergence analysis of both the C-gene and the complete genome (Fig. 2).

Among the strains analyzed, 8 strains were found to contain substitutions within the S-gene encoding for the HBsAg, that have been associated with vaccine escape (Table 3). No substitutions associated with antiviral resistance and no marked insertions or deletions were identified.

3.4. Recombinant HBV C/D strain

Recombinant strain 17-126, was isolated in 2017 from a female born in the Netherlands with an acute HBV infection.

Analysis of the S-gene indicates the strain 17-126 to be most similar to genotype D strains in our study (divergence of 1.20 %, while

Table 3
Number of strains with substitutions associated with vaccine escape (VEM).

	Genotype		
Infection status	A	B	D
Acute	P120T (n = 3)	T126A (n = 1) T126S (n = 1)	
Chronic	P120T (n = 1)		Q129H (n = 1)
Unknown			D144E (n = 1)

the C-gene showed a lowest divergence of 1.90 % to genotype C. Divergence based on the complete genome was the lowest compared to genotype C (4.4 %). Bootscan analysis indicates a breakpoint at the preS2- and S-gene junction and at the 5′end of the S-gene (Fig. 3).

4. Discussion

To date, 10 HBV genotypes are known with each its own genetic, epidemiological and clinical patterns and accurate characterization and analysis of the different circulating strains is essential for surveillance and to gain insight in transmission patterns.

In this study, we developed a complete genome assay for HBV genotyping which shows the value of complete genome analysis versus the use of only S-, and C-gene analysis for surveillance. The assay enabled successful sequencing of 79.8 % of the samples tested, with a lower limit of detection of 500 IU/ml. Failure of the assay was mostly due to low or no DNA present as indicated by the high ct values.

In concordance with the geographical distribution in Europe, we found a high prevalence of genotype A followed by genotype D [7, 8]. HBV genotypes B, C, E, F, and H were detected as minority genotypes. The higher genetic variation of the C-gene compared to the S-gene is in line with what was previously reported [2]. While complete genome analysis also shows a statitically higher genetic variation compared to the S-gene for all genotypes, it also showed a statical difference compared to the C-gene for genotypes A through E. This shows the added value of complete genome analysis over separate S- and C- gene analyses. Difference between the mean within group divergence of different genotypes was further observed and was lowest for genotype F, followed by genotype A for all three analyses. The low divergence concurs with the endemic circulation of genotype A in the Netherlands. In addition, the low divergence of genotype F in the 3 analyses, confirms previous observations of this type in the Netherlands [2] where it was suggested to be continued endemic circulation, most probably with a source from the Americas, where the type is endemic. The higher divergence seen for other genotypes, in combination with the low numbers due to low circulation, suggest circulation of these types most likely to be due to importation [27].

It should be noted that the divergence calculations are predominated by strains from acute infected cases in the Netherlands where divergence is generally low compared to chronic cases [28] and comparison with other studies are limited due to differences in populations studied.

The high variation in the C-gene enables transmission analysis but is inaccurate for typing, in particular for the B/C and D/E genotypes due to intragenotypic recombination between these strains in the past [15]. As the variation in the S-gene used for typing is lower, genetic clusters used to analyze transmission patterns, in particular for gA, are indistinguishable. However, complete genome analysis showed a high genetic resolution to enable transmission analysis, and simultaneously enables HBV typing, thus providing accurate characterization and analysis of the different circulating strains for public health surveillance.

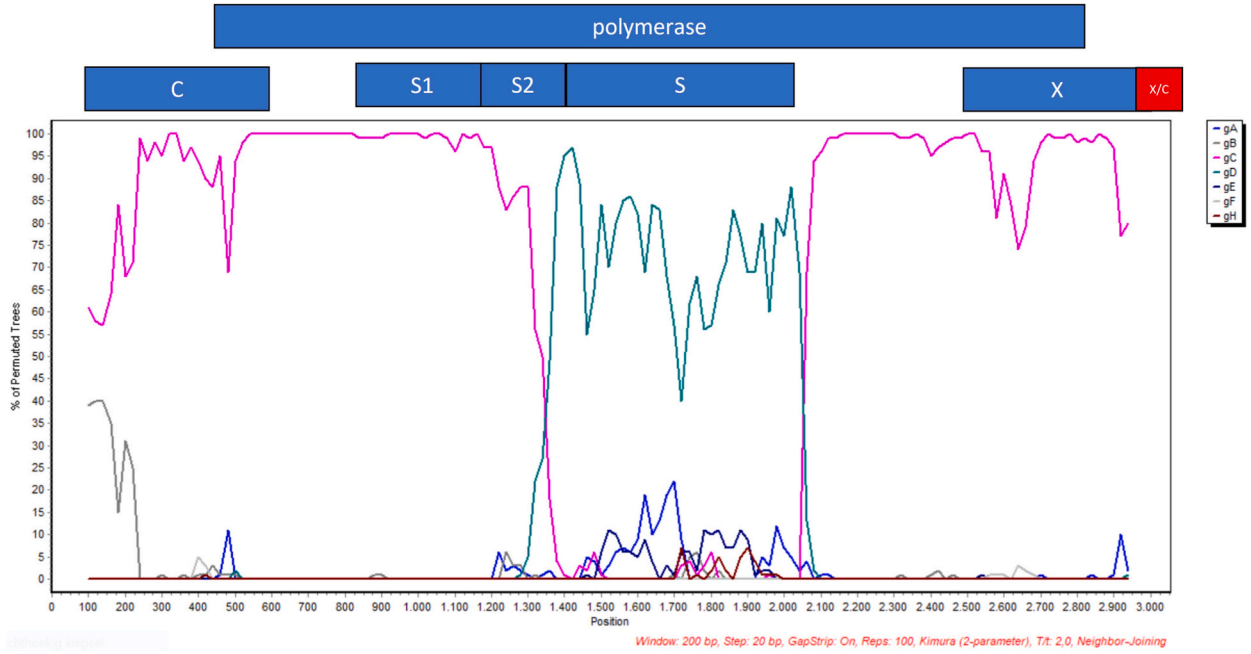


Fig. 3. Bootscan plot of recombinant strain (417-126/gCD) against study sequences representing genotype A-H (417-136/gA, 418-96/gB, 418-108/gC, 418-109/gD, 419-73/gE, 418-99/gF, and 418-61/gH). Each point represents the percentage identity within a sliding window 200 bp wide, centered on the position plotted, with a step size of 20 bp. Positions containing gaps were excluded from the comparison by gap stripping, and Kimura-2-parameter correction was applied. The genome covered by the complete genome assay is indicated in blue. Simplot and Bootscan plots were generated using software version 3.5.1.

In this study, we identified a recombinant gC/D strain by analysis of the complete genome along with partial analysis of the genes. In concordance with criteria stated by Pourkarim et al. [17], the recombinant strain was excluded from phylogenetic analysis and analyzed separately. S-gene analysis alone would suggest the type to be a gD. However C-gene analysis and complete genome analysis suggest the strain to be recombinant strain with breakpoints flanking the S-gene. The Dutch strain showed similar genetic characteristics as two recombinant C/D strains that were isolated in China [29].

In conclusion, these data stress the importance of analyzing complete genomes for accurate characterization and transmission analysis of circulating strains for public health surveillance. In the current study, Sanger sequencing was used to generate complete genomes. The assay uses only two primer pairs for the generation of the complete genome-amplicon and 8 sequencing primers, which is less labor intensive than other complete genome assays using multiple primer pairs to generate the complete genome-amplicon and multiple sequencing primers [30,31]. Advances in sequence technology, such as Nanopore and Ilumina will further enable analysis of the viral genetic variability within HBV infected cases [32,33].

Funding

This study was supported by the Ministry of Health, Welfare and Sport, The Netherlands. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Ethical statement

Informed consent was not required for this study because all cases positive for Hepatitis B virus are mandatorily notified by the Municipal Health Centers and typed according to Dutch Public Health law (WPG), article 25(5).

During the preparation of this work the author(s) did not use any AI tools.

CRediT authorship contribution statement

Jeroen Cremer: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Francoise van Heiningen:** Validation, Writing – review & editing. **Irene Veldhuijzen:** Supervision, Validation, Writing – review & editing. **Kimberley Benschop:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Kahren van Eer for critically reading the manuscript and the laboratory staff of Infectious diseases Diagnosis and Screening of the Center for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands, for typing all the serum samples.

References

[1] A. Schweitzer, J. Horn, R.T. Mikolajczyk, G. Krause, J.J. Ott, Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013, Lancet 386 (10003) (2015) 1546–1555, https://doi.org/10.1016/S0140-6736(15)61412-X.

[2] H.J. Boot, J. Cremer, F.D. Koedijk, W.M. van Ballegooijen, E.L. Op de Coul, Improved tracing of hepatitis B virus transmission chains by phylogenetic analysis based on C region sequences, J. Med. Virol. 80 (2) (2008) 233–241, https://doi.org/10.1002/jmv.21080.

[3] M.R. Pourkarim, S. Amini-Bavil-Olyaee, F. Kurbanov, M. Van Ranst, F. Tacke, Molecular identification of hepatitis B virus genotypes/subgenotypes: revised classification hurdles and updated resolutions, World J. Gastroenterol. : WJG 20 (23) (2014) 7152–7168, https://doi.org/10.3748/wjg.v20.i23.7152.

[4] A. Kramvis, M. Kew, G. Francois, Hepatitis B virus genotypes, Vaccine 23 (19) (2005) 2409–2423, https://doi.org/10.1016/j.vaccine.2004.10.045.

[5] H. Okamoto, F. Tsuda, H. Sakugawa, R.I. Sastrosoewignjo, M. Imai, Y. Miyakawa, M. Mayumi, Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes, The Journal of general virology 69 (Pt 10) (1988) 2575–2583, https://doi.org/10.1099/0022-1317-69-10-2575.

[6] S. Schaefer, Hepatitis B virus taxonomy and hepatitis B virus genotypes, World J. Gastroenterol. : WJG 13 (1) (2007) 14–21, https://doi.org/10.3748/wjg.v13.i1.14.

[7] R. van Houdt, G.J. Sonder, N.H. Dukers, L.P. Bovee, A. van den Hoek, R.A. Coutinho, S.M. Bruisten, Impact of a targeted hepatitis B vaccination program in Amsterdam, The Netherlands, Vaccine 25 (14) (2007) 2698–2705, https://doi.org/10.1016/j.vaccine.2006.06.058.

[8] S. Hahne, R. van Houdt, F. Koedijk, M. van Ballegooijen, J. Cremer, S. Bruisten, R. Coutinho, H. Boot, Selective hepatitis B virus vaccination has reduced hepatitis B virus transmission in The Netherlands, PLoS One 8 (7) (2013), e67866, https://doi.org/10.1371/journal.pone.0067866.

[9] N. Akuta, H. Kumada, Influence of hepatitis B virus genotypes on the response to antiviral therapies, J. Antimicrob. Chemother. 55 (2) (2005) 139–142, https://doi.org/10.1093/jac/dkh533.

[10] T. Guettouche, H.J. Hnatyszyn, Chronic hepatitis B and viral genotype: the clinical significance of determining HBV genotypes, Antivir. Ther. 10 (5) (2005) 593–604.

[11] P.K. Chandra, A. Biswas, S. Datta, A. Banerjee, R. Panigrahi, S. Chakrabarti, B.K. De, R. Chakravarty, Subgenotypes of hepatitis B virus genotype D (D1, D2, D3 and D5) in India: differential pattern of mutations, liver injury and occult HBV infection, J. Viral Hepat. 16 (10) (2009) 749–756, https://doi.org/10.1111/j.1365-2893.2009.01129.x.

[12] W. Shi, Z. Zhang, C. Ling, W. Zheng, C. Zhu, M.J. Carr, D.G. Higgins, Hepatitis B virus subgenotyping: history, effects of recombination, misclassifications, and corrections, Infect. Genet. Evol. : journal of molecular epidemiology and evolutionary genetics in infectious diseases 16 (2013) 355–361, https://doi.org/10.1016/j.meegid.2013.03.021.
[13] N.M. Araujo, Hepatitis B virus intergenotypic recombinants worldwide: an overview, Infect. Genet. Evol. : journal of molecular epidemiology and evolutionary genetics in infectious diseases 36 (2015) 500–510, https://doi.org/10.1016/j.meegid.2015.08.024.
[14] L. Lopez, D. Flichman, L. Mojsiejczuk, M.V. Gonzalez, R. Uriarte, R. Campos, J. Cristina, L. Garcia-Aguirre, Genetic variability of hepatitis B virus in Uruguay: D/F, A/F genotype recombinants, Arch. Virol. 160 (9) (2015) 2209–2217, https://doi.org/10.1007/s00705-015-2477-0.
[15] P. Simmonds, S. Midgley, Recombination in the genesis and evolution of hepatitis B virus genotypes, J. Virol. 79 (24) (2005) 15467–15476, https://doi.org/10.1128/JVI.79.24.15467-15476.2005.
[16] C.M. Martin, J.A. Welge, J.T. Blackard, Hepatitis B virus (HBV) X gene diversity and evidence of recombination in HBV/HIV co-infected persons, J. Med. Virol. 83 (7) (2011) 1142–1150, https://doi.org/10.1002/jmv.22090.
[17] M.R. Pourkarim, S. Amini-Bavil-Olyaee, P. Lemey, P. Maes, M. Van Ranst, Are hepatitis B virus "subgenotypes" defined accurately? J. Clin. Virol. : the official publication of the Pan American Society for Clinical Virology 47 (4) (2010) 356–360, https://doi.org/10.1016/j.jcv.2010.01.015.
[18] F. Sugauchi, E. Orito, T. Ichida, H. Kato, H. Sakugawa, S. Kakumu, T. Ishida, A. Chutaputti, C.L. Lai, R. Ueda, et al., Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene, J. Virol. 76 (12) (2002) 5985–5992, https://doi.org/10.1128/jvi.76.12.5985-5992.2002.
[19] T. Sakamoto, Y. Tanaka, J. Simonetti, C. Osiowy, M.L. Borresen, A. Koch, F. Kurbanov, M. Sugiyama, G.Y. Minuk, B.J. McMahon, et al., Classification of hepatitis B virus genotype B into 2 major types based on characterization of a novel subgenotype in Arctic indigenous populations, J. Infect. Dis. 196 (10) (2007) 1487–1492, https://doi.org/10.1086/523111.
[20] P.S.S.E. Simmonds, A nucleotide and amino acid sequence analysis platform, BMC Res. Notes 5 (2012) 50, https://doi.org/10.1186/1756-0500-5-50.
[21] A. Bartholomeusz, S. Schaefer, Hepatitis B virus genotypes: comparison of genotyping methods, Rev. Med. Virol. 14 (1) (2004) 3–16, https://doi.org/10.1002/rmv.400.
[22] J. Hadfield, C. Megill, S.M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, R.A. Neher, Nextstrain: real-time tracking of pathogen evolution, Bioinformatics 34 (23) (2018) 4121–4123, https://doi.org/10.1093/bioinformatics/bty407.
[23] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 (14) (2002) 3059–3066, https://doi.org/10.1093/nar/gkf436.
[24] L.T. Nguyen, H.A. Schmidt, A. von Haeseler, B.Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, Mol. Biol. Evol. 32 (1) (2015) 268–274, https://doi.org/10.1093/molbev/msu300.
[25] P. Sagulenko, V. Puller, R.A. Neher, TreeTime: maximum-likelihood phylodynamic analysis, Virus Evol 4 (1) (2018) vex042, https://doi.org/10.1093/ve/vex042.
[26] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (7) (2016) 1870–1874, https://doi.org/10.1093/molbev/msw054.
[27] M. Kretzschmar, G.A. de Wit, L.J. Smits, M.J. van de Laar, Vaccination against hepatitis B in low endemic countries, Epidemiol. Infect. 128 (2) (2002) 229–244.
[28] J. Cremer, S.H.I. Hofstraat, F. van Heiningen, I.K. Veldhuijzen, B.H.B. van Benthem, K.S.M. Benschop, Genetic variation of hepatitis B surface antigen among acute and chronic hepatitis B virus infections in The Netherlands, J. Med. Virol. 90 (10) (2018) 1576–1585, https://doi.org/10.1002/jmv.25232. From NLM Medline.
[29] Z. Wang, Z. Liu, G. Zeng, S. Wen, Y. Qi, S. Ma, N.V. Naoumov, J. Hou, A new intertype recombinant between genotypes C and D of hepatitis B virus identified in China, The Journal of general virology 86 (Pt 4) (2005) 985–990, https://doi.org/10.1099/vir.0.80771-0.
[30] S. Gunther, B.C. Li, S. Miska, D.H. Kruger, H. Meisel, H. Will, 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 (9) (1995) 5437–5444, https://doi.org/10.1128/JVI.69.9.5437-5444.1995. From NLM Medline.
[31] R.W. Lieshout-Krikke, M.W. Molenaar-de Backer, P. van Swieten, H.L. Zaaijer, Surface antigen-negative hepatitis B virus infection in Dutch blood donors, Eur. J. Clin. Microbiol. Infect. Dis. : official publication of the European Society of Clinical Microbiology 33 (1) (2014) 69–77, https://doi.org/10.1007/s10096-013-1930-9. From LM Medline.N.
[32] S. Ghosh, A. Chakraborty, S. Banerjee, Persistence of hepatitis B virus infection: a multi-faceted player for hepatocarcinogenesis, Front. Microbiol. 12 (2021), 678537, https://doi.org/10.3389/fmicb.2021.678537. From NLM PubMed-not-MEDLINE.
[33] V. Sauvage, L. Boizeau, D. Candotti, M. Vandenbogaert, A. Servant-Delmas, V. Caro, S. Laperche, Early MinION nanopore single-molecule sequencing technology enables the characterization of hepatitis B virus genetic complexity in clinical samples, PLoS One 13 (3) (2018), e0194366, https://doi.org/10.1371/journal.pone.0194366. From NLM Medline.