TITLE:

Evolutional transition of HBV genome during the persistent infection determined by single-molecule real-time sequencing

AUTHOR(S):

Arasawa, Soichi; Takeda, Haruhiko; Takai, Atsushi; Iguchi, Eriko; Eso, Yuji; Shimizu, Takahiro; Takahashi, Ken; ... Ueda, Yoshihide; Marusawa, Hiroyuki; Seno, Hiroshi

CITATION:

Arasawa, Soichi ...[et al]. Evolutional transition of HBV genome during the persistent infection determined by single-molecule real-time sequencing. Hepatology Communications 2023, 7(3): e0047.

ISSUE DATE:

2023-03

URL:

http://hdl.handle.net/2433/284653

RIGHT:

Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Association for the Study of Liver Diseases.; This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Received: 30 November 2022 | Accepted: 20 December 2022

DOI: 10.1097/HC9.0000000000000047

ORIGINAL ARTICLE

OPEN

Evolutional transition of HBV genome during the persistent infection determined by single-molecule real-time sequencing

Soichi Arasawa[1] | **Haruhiko Takeda**[1] | **Atsushi Takai**[1] | **Eriko Iguchi**[1] | **Yuji Eso**[1] | **Takahiro Shimizu**[1] | **Ken Takahashi**[1] | **Taiki Yamashita**[2] | **Yoshihide Ueda**[3] | **Hiroyuki Marusawa**[4] | **Hiroshi Seno**[1]

[1]Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[2]Department of Primary Care and Emergency Medicine, Graduate School of Medicine, Chiba University, Chiba, Japan

[3]Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kobe University, Kobe, Japan

[4]Department of Gastroenterology and Hepatology, Osaka Red Cross Hospital, Osaka, Japan

Correspondence
Atsushi Takai, Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, 54 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan.
E-mail: atsushit@kuhp.kyoto-u.ac.jp

Abstract

Background. Although HBV infection is a serious health issue worldwide, the landscape of HBV genome dynamics in the host has not yet been clarified. This study aimed to determine the continuous genome sequence of each HBV clone using a single-molecule real-time sequencing platform, and clarify the dynamics of structural abnormalities during persistent HBV infection without antiviral therapy.

Patients and Methods. Twenty-five serum specimens were collected from 10 untreated HBV-infected patients. Continuous whole-genome sequencing of each clone was performed using a PacBio Sequel sequencer; the relationship between genomic variations and clinical information was analyzed. The diversity and phylogeny of the viral clones with structural variations were also analyzed.

Results. The whole-genome sequences of 797,352 HBV clones were determined. The deletion was the most common structural abnormality and concentrated in the preS/*S* and *C* regions. Hepatitis B e antibody (anti-HBe)-negative samples or samples with high alanine aminotransferase levels have significantly diverse deletions than anti-HBe-positive samples or samples with low alanine aminotransferase levels. Phylogenetic analysis demonstrated that various defective and full-length clones evolve independently and form diverse viral populations.

Conclusions. Single-molecule real-time long-read sequencing revealed the dynamics of genomic quasispecies during the natural course of chronic

---

Abbreviations: ALT, alanine aminotransferase; CCS, circular consensus sequence; FL, full-length, NGS, next-generation sequencing; RT, reverse transcriptase; SV, structural variation; SMRT, single-molecule real-time; SNV, single nucleotide variant; TP, terminal protein.

Soichi Arasawa and Haruhiko Takeda contributed equally to this work.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.hepcommjournal.com

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.
Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Association for the Study of Liver Diseases.

HBV infections. Defective viral clones are prone to emerge under the condition of active hepatitis, and several types of defective variants can evolve independently of the viral clones with the full-length genome.

INTRODUCTION

HBV is a partially double-stranded anthropotropic DNA virus belonging to the Orthohepadnavirus genus of the Hepadnaviridae family in the Riboviria realm.[1] HBV has an ~3.2 kb-long open, circular genome containing 4 overlapping open reading frames. It has a unique life history, with persistently present covalently closed circular DNA as the genetic template, for sustained viral replication through reverse transcription of pregenomic RNA in host cells.[2,3] As with other viruses in Hepadnaviridae, HBV causes hepatotropic infections in its natural hosts.[4] According to the World Health Organization, ~296 million people were positive for HBV in 2019. HBV is a major cause of severe health problems such as acute hepatitis, chronic hepatitis, liver cirrhosis, and HCC.[5]

Paleovirological studies have revealed the evolutionary history of Hepadnaviridae in an estimated > 80 million years since the emergence of their ancestor.[6] Gilbert and Feschotte[7] identified a dozen endogenous duck HBV-like sequences in the zebra finch genome using a BLAST-based genomic analysis. Presence/absence analyses performed by Suh et al [9] revealed that the long-term substitution rate of Hepadnaviridae between the Mesozoic and Miocene periods was in the order of 10^{-8} per site per year, which is ~10^{-3} times slower than the short-time substitution rate of human HBV in several decades.[8] Suh et al[9] also reported that the amino-acid sequences of extant Hepadnaviridae genomes are highly conserved in several regions, such as the reverse transcriptase, *S*, *C*, terminal protein (TP), and ribonuclease *H* (RNaseH) regions, at a rate of 12.2%–27.1%, as compared with the hypothetical ancestral sequence in the Mesozoic era.

In terms of genomic size, previous studies using conventional methods such as Southern blot, PCR-based electrophoresis, and direct sequencing have revealed the presence of both interspecific and intraspecific variation among Hepadnaviridae viruses. For example, duck HBV has a smaller genome (with a total length of 3021–3027 bases), whereas the woodchuck hepatitis virus has a larger genome (3308–3320 bases) than the human HBV (3182–3284 bases).[6] Conventional studies have also revealed genomic size variants described as defective HBV particles derived from spliced pregenomic RNA, even in a single patient with chronic HBV infection.[10–14] Recently, Suzuki et al[15] used short-read ultra-deep next-generation sequencing to report that a 27-base-long deletion in the *preS2* gene was associated with the higher serological activity of HBV and advanced liver fibrosis. Betz-Stablein et al[16] used single-molecule real-time (SMRT) sequencing to detect various spliced variants of HBV with 1 kb or larger deletions, in a case of chronic HBV infection in a patient who underwent liver transplantation and > 10 years of antiviral treatment. However, the evolutionary dynamics of genomic structural variant-based haplotypes, including defective HBV genomes in each case of chronic HBV infection in natural, treatment-free internal environments, and the relationship between those haplotypes and intraspecific variants of human HBV are still unclear. To clarify the dynamics of HBV genomic quasispecies, we performed SMRT sequencing-based analyses of the HBV genome derived from cases of chronic infection without any treatment.

PATIENTS AND METHODS

Patients and sample collection

HBV-seropositive patients who received clinical observation at Kyoto University Hospital and Osaka Red Cross Hospital between April 2001 and March 2020 were enrolled in this study. The inclusion criteria were: consistently seropositive for HBsAg serum HBV-DNA titer higher than 4.23 LogIU/mL (10^5 copies/mL) for at least 2 time-points; free from antiviral medications or interferon treatment. The exclusion criteria were: under 20 years old; coinfection with HCV, HIV, or human T-lymphotropic virus; uncontrolled malignancies; liver comorbidities such as alcohol-associated liver disease, nonalcoholic steatohepatitis, autoimmune hepatitis, primary biliary cholangitis, primary sclerosing cholangitis, immunoglobulin G4-related sclerosing cholangitis, or Wilson disease. Two to 5 serum samples were sequentially collected during the observation period. All samples were surpluses of sera drawn for clinical purposes. The collected samples were stored at −80°C.

This study conformed to the provisions of the Declaration of Helsinki. The study protocol (R2594) was approved by the ethics committee of Kyoto University and the clinical samples were obtained with written informed consent or based on an opt-out method of consent from all participants.

Single-molecule real-time sequencing

Viral DNA was extracted from the sera and the almost full-length viral genomic DNA was amplified from it using PCR. Long-read sequencing spanning 3000 contiguous bases of HBV-DNA was performed using the SMRTseq platform with PacBio Sequel, following the protocol from Pacific Biosciences (Menlo Park). To improve the accuracy of the sequenced reads, circular consensus sequence 2 (CCS2) reads were generated; of these, ≥5-pass CCS2 reads confirmed to have quite low error rates were used for genetic analysis.[17,18] Single nucleotide variants (SNVs) and structural variations (SVs) were detected at the single-clone resolution, and these genetic data were phylogenetically analyzed. The details are described in the “supplementary methods” section, http://links.lww.com/HC9/A238.

Conventional short-read next-generation sequencing

Ultra-deep sequencing with multiplexed tags was performed using an Ion Proton Sequencer (Thermo Fisher Scientific). The details are described in the “supplementary methods” section, http://links.lww.com/HC9/A238.

Statistical analyses

Statistical analyses were performed using R v4.0.1 (The R Foundation for Statistical Computing). Categorical values were analyzed using Fisher exact test, whereas continuous values were analyzed using the Mann-Whitney *U* test. Differences were considered statistically significant at $p < 0.05$.

Sequence data availability

The sequencing results of this study were submitted to the DNA Data Bank of Japan Sequence Read Archive (https://www.ddbj.nig.ac.jp/dra/index-e.html) under accession number DRA014988.

PCR and library preparation, long-read sequencing and the generation of CCS, short-read ultra-deep sequencing, detection of genomic SVs, validation of representative SVs, and phylogenetic analysis are described in the “supplementary materials and methods”, http://links.lww.com/HC9/A238.

RESULTS

Sequential sera sample collection from chronic hepatitis B patients

We collected 25 serum samples from 10 cases, 5 male patients and 5 female patients with a median age of 39 years (range = 19–84 years), at the first time-point of the observation period. Five cases had genotype *B*, whereas the others had genotype *C*. The other clinical characteristics are detailed in Supplementary Table 1 (http://links.lww.com/HC9/A107). None of the patients were administered any anti-HBV agents, including interferon and nucleotide analogs. Mean observational period in the current study was 4.1 years (range = 2.0–9.3 years).

The patients were classified into 3 clinical groups based on their immunologic status. The first group consisted of 3 patients (cases #01, #02, and #03) who were consistently positive for HBeAg and negative for hepatitis B e antibody (anti-HBe) during the observation period. Serum alanine aminotransferase (ALT) levels (IU/mL) were within normal ranges. The HBV-DNA levels (logIU/mL) in the serum samples from cases #01 and #03 persisted at high levels, whereas those in serum samples from case #02 changed dynamically throughout the observation period (Figure 1A). The second group included 4 cases (cases #04, #05, #06, and #07), in which the positivity of HBeAg and anti-HBe changed during the observation period. This group showed dynamic chronological changes in the serum HBV-DNA and ALT levels. Cases #05 and #06 demonstrated a severe immunologic response against the viruses. In each case, anti-HBe seemed during the clinical course, and cases #06 and #07 achieved seroconversion (Figure 1B). The last group included the remaining 3 cases (cases #08, #09, and #10) that were negative for HBeAg and positive for anti-HBe throughout the clinical course. Serum ALT levels were consistently suppressed during the observation period (Figure 1C). Serum samples were sequentially collected at 2–5 time points from each patient and subjected to sequencing analysis.

Nearly full-length determination of continuous HBV genome at viral clone resolution

We performed SMRT sequencing of each sample to examine variations in the viral genome sequence. Figure 2 depicts the workflow for the sequencing analysis. The HBV-DNA extracted from sequentially collected serum samples was amplified and processed for SMRT sequencing using the PacBio Sequel platform. CCS2 reads were computationally generated, and ≥5-pass CCS2 reads were selected to exclude low-quality sequence reads from further analysis.

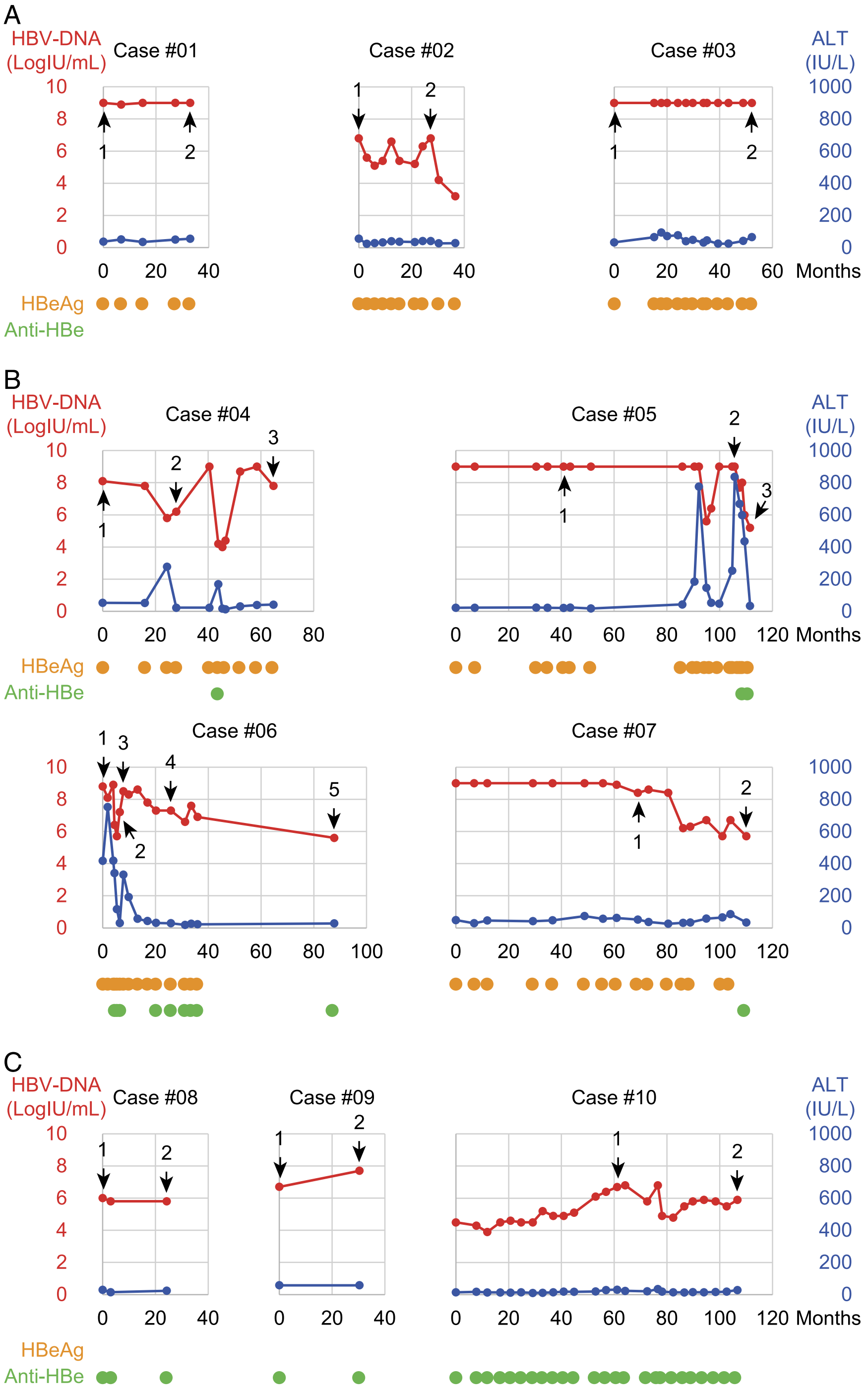

FIGURE 1 Clinical courses and sample collection. The clinical courses of all cases included in this study are shown. (A) Cases negative for anti-HBe throughout the observational period (cases #01–#03). (B) Cases that experienced seroconversion; anti-HBe was negative at the first time-point and then became positive at least one time-point during the observational period (cases #04–#07). (C) Cases positive for anti-HBe throughout the observational period (cases #08–#10). The *x*-axis represents the months after the first visit to the hospital in each case. The left-side *y*-axis represents the HBV-DNA level (LogIU/mL), whereas the right-side *y*-axis represents the serum ALT level (IU/mL). The red and blue broken lines represent the HBV-DNA and ALT levels in the graphs, respectively. The orange and green marks at the bottom of each graph represent a sample being positive for HBeAg and anti-HBe, respectively. The black arrows represent the time points, at which the serum samples were collected. Abbreviation: ALT indicates alanine aminotransferase.

To determine the accuracy of $\geq$5-pass CCS2 reads from HBV clones, we first examined the sequencing error rate of CCS2 reads derived from HBV-containing plasmids, which are generally regarded as monoclonal. In the duplicate experiments, we confirmed extremely low error rates of $\geq$5-pass CCS2 reads derived from

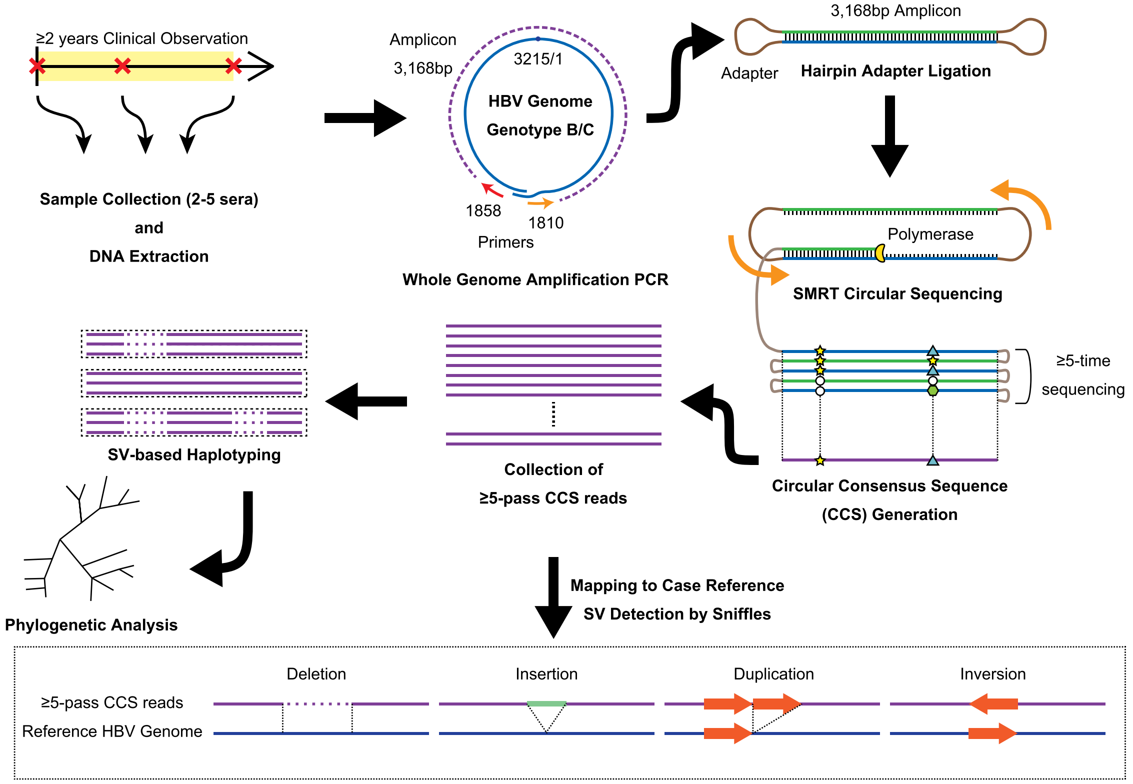
≥2 years Clinical Observation
Sample Collection (2-5 sera)
and
DNA Extraction
Amplicon
3,168bp
3215/1
HBV Genome
Genotype B/C
1858
1810
Primers
Whole Genome Amplification PCR
3,168bp Amplicon
Adapter
Hairpin Adapter Ligation
Polymerase
SMRT Circular Sequencing
≥5-time
sequencing
Circular Consensus Sequence
(CCS) Generation
Collection of
≥5-pass CCS reads
SV-based Haplotyping
Phylogenetic Analysis
Mapping to Case Reference
SV Detection by Sniffles
Deletion
Insertion
Duplication
Inversion
≥5-pass CCS reads
Reference HBV Genome

FIGURE 2 Workflow of the PacBio sequencing analysis. This flow demonstrates the protocol for the experiments. Multiple serum samples were chronologically collected throughout the observational period of more than 2 years. Nearly the whole-genome region of HBV was amplified, ligated with hairpin adapters for PacBio sequencing, and subjected to SMRT CCS2 with the PacBio Sequel sequencer. High-quality contiguous sequence reads (≥5-pass CCS2 reads) were generated, and bioinformatic analyses were conducted. CCS2 reads mapping was carried out using the NGMLR algorithm, for the detection of structural variants with Sniffles software. The phylogenetic trees were constructed based on the CCS2 reads sorted according to haplotypes using MEGA X software. Abbreviation: CCS indicates circular consensus sequence; SMRT, single-molecule real-time.

HBV-containing plasmids; 0.00% for SVs (Supplementary Table 2, http://links.lww.com/HC9/A107). Based on these data, we applied well-qualified >5-pass CCS2 reads from clinical HBV samples for further analysis.

In total, we determined 797,352 ≥5-pass CCS2 reads derived from the HBV genome from 25 samples (median 31,894 reads per sample) (Supplementary Table 3, http://links.lww.com/HC9/A107). Almost all reads had sequence lengths of 3168 bp, suggesting that near full-length contiguous HBV genome sequences from each HBV clone that simultaneously infected the host were determined using the SMRT sequencing. From all serum samples (25 out of 25), we successfully achieved near full-length continuous determination of HBV genome sequences of more than 9000 independent HBV clones per serum sample, which form viral quasispecies in each host.

Deletion is the predominant type of structural variation

Figure 3 shows the entire genome mapping of SMRT CCS2 reads obtained from each serum sample of the representative case (case #04). In contrast to short-read sequencing, the sequence depth showed a uniform distribution, because the CCS2 reads were contiguous whole-genome sequences of HBV clones. However, the sequence depth of specific genomic regions is shallow if there exists a subclonal population of HBV with genomic deletion, the so-called defective clone, among the HBV quasispecies. Undulation of the sequence depth of conventional next-generation sequencing hinders the detection of these deletions, whereas the plateaued depth of SMRT sequencing reveals the crevasses corresponding to the deletions and their frequencies (Figure 3).

Using long-read sequencing-specific bioinformatics tools for genome mapping and SV detection, we detected a total of 1736 SVs among all ≥5-pass CCS2 reads from the 25 samples (Supplementary Table 4, http://links.lww.com/HC9/A107). We found that 431 SVs were with a variant frequency of >0.1% among all sequence reads in each sample; of these, 407 SVs (94.4%) were deletions, 22/431 SVs (5.1%) were insertions, and the remaining 2 SVs were duplications (Figure 4A and Supplementary Table 5, http://links.lww.com/HC9/A107). Most serum samples contained viral clones with SVs of various lengths.

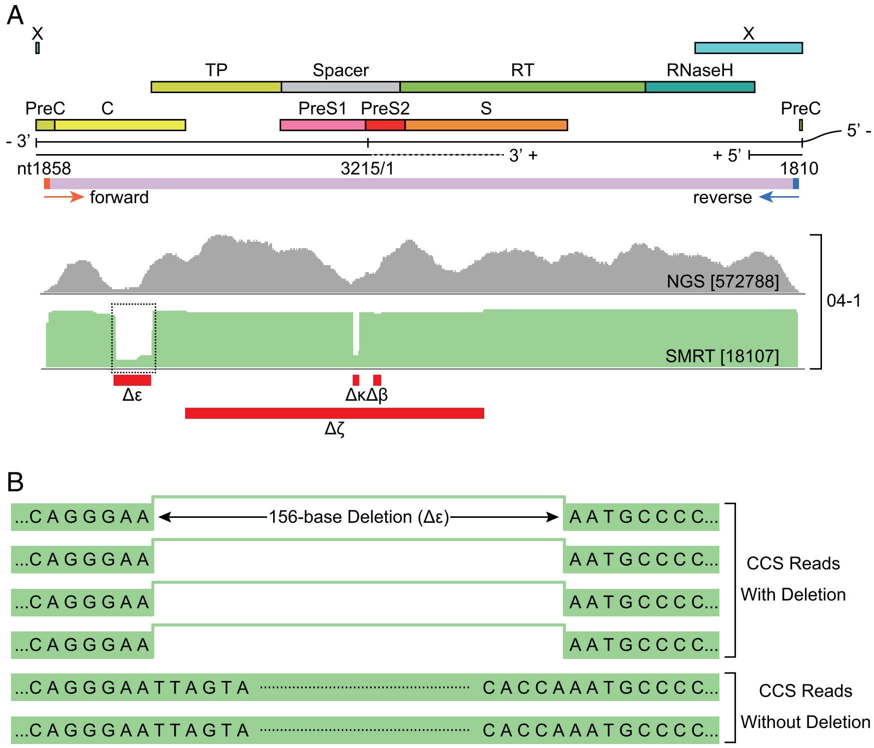

FIGURE 3 Mapping images of the entire HBV genome sequence, with genetic deletions determined using SMRT sequencing. Schematic images of the genetic regions of the reference sequence of the incomplete double-strand DNA of HBV are shown at the top of this panel. (A) Comparison of the mapped reads of conventional NGS (gray) and SMRT sequencing (green) of sample 04-1. (B) The magnified image of the mapped reads is surrounded by a dotted line. A scheme of g.2150_2305del (Δε) is shown. Abbreviation: NGS indicates next-generation sequencing; RT, reverse transcriptase; SMRT, single-molecule real-time; TP, terminal protein.

We focused on HBV clones, in which SVs were detected at a high frequency. Among these SVs, in total 17 and in actual 13 loci of deletions were detected with a frequency of ≥10% (Figure 4B and Supplementary Table 6, http://links.lww.com/HC9/A107). Two large deletions with a length >1 kb affecting multiple genes of the HBV genome were detected (Δδ and Δζ), one of which has been identified as the SP1 spliced variant in previous reports (Δζ).[11,13,14,16,19–26] The remaining 11 deletions were shorter than 200 bases, affecting 1 or 2 genes. Five deletions affected preS1/spacer (Δη, Δη′, Δθ, Δι, and Δκ), whereas 2 each affected preS2/spacer (Δα and Δβ), *X* gene (Δγ and Δγ′), and *C* gene (Δε and Δε′). The prevalence of in-frame deletions was significantly higher in these major deletions than in SVs with frequencies <10% (Supplementary Table 7, http://links.lww.com/HC9/A107).

To validate the deletions detected using PacBio sequencing, we conducted short-read high-throughput sequencing and agarose gel electrophoresis for the representative deletions. As shown in Figure 4C and Supplementary Figure 1 (http://links.lww.com/HC9/A102), short-read fragments with specific breakpoints were mapped to the case-specific reference genome sequence, thereby confirming the existence of HBV-DNA with specific deletions. In addition, we validated 2 representative deletions (Δζ and Δι) using breakpoint-specific PCR conducted on samples 03-2 and 04-3, respectively (Figure 4D).

Anti-HBV immunologic status correlated with the frequency of structural variations and precore hotspot mutations

We examined the association between the number of SVs and hepatitis activity (Supplementary Table 8, http://links.lww.com/HC9/A107). Interestingly, the total numbers of SVs/deletions were significantly larger in anti-HBe-negative and HBeAg-positive samples, respectively (Figure. 5A, B). Furthermore, there were more SVs/deletions in the sera with higher ALT levels than in those with lower ALT levels (Figure 5C). In contrast, there was no significant difference between high and low HBV-DNA titers (Supplementary Figure 2A, http://links.lww.com/HC9/A103). Region-associated analyses of deletions revealed that the preS1, *S*, *C*, TP, spacer, and reverse transcriptase regions had higher numbers of deletions in anti-HBe-negative samples than in anti-HBe-positive samples (Figure 5D).

As mutations in the precore or core promoter region are known to be associated with HBeAg seroconversion, we examined the correlation between the allele frequency of hotspot mutations in precore or core promoter and HBeAg status. As reported, the G1896A mutation in the precore region and T1762A mutation in the core promoter region were more frequently identified in HBeAg-negative samples than in

京都大学 KYOTO UNIVERSITY
A Self-archived copy in Kyoto University Research Information Repository https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

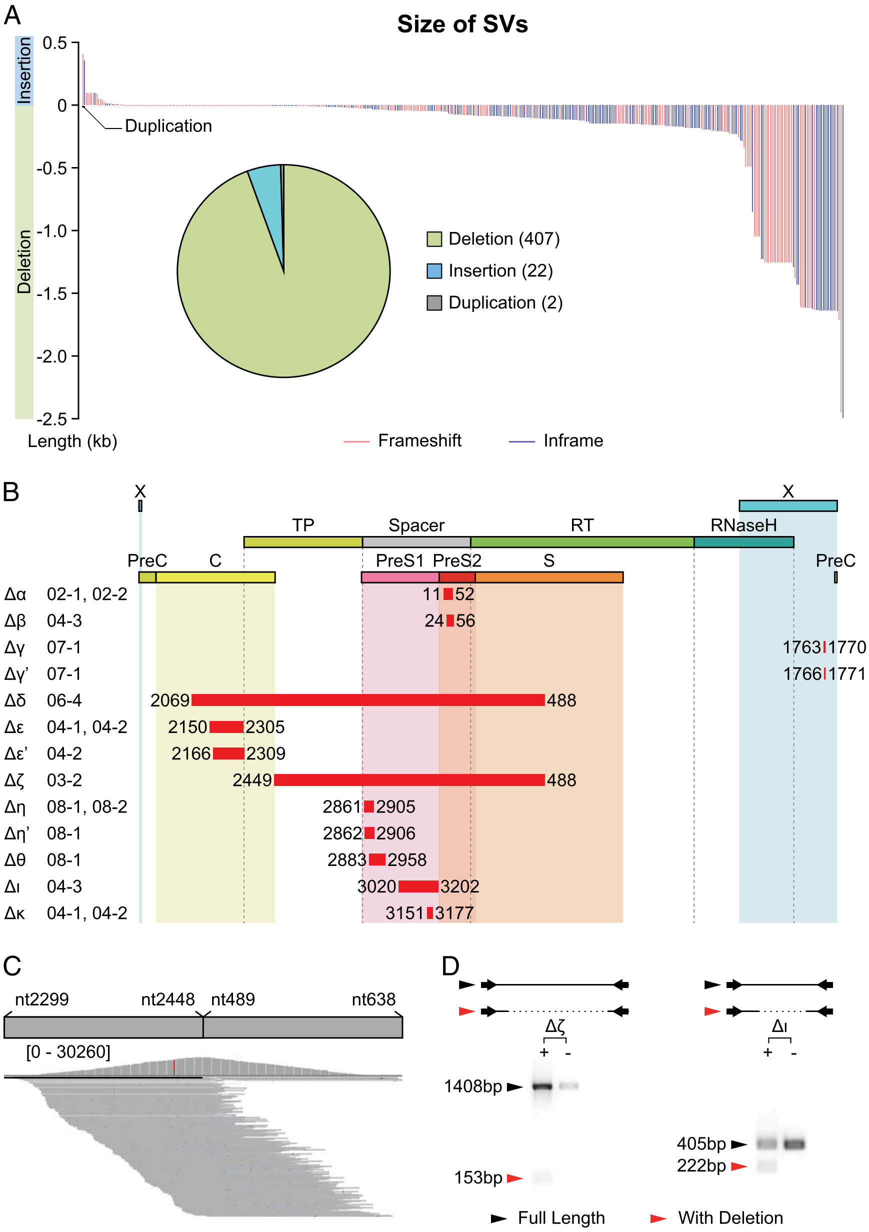

FIGURE 4 Characteristics of SVs. (A) Overview of SVs with frequencies > 0.001. A pie chart showing the sorting of the SVs. Identical SVs detected in 2 or more samples were counted individually. The waterfall plot represents the distribution of the SV sizes. Positive and negative numbers indicate insertions and deletions, respectively. Red and blue bars represent frameshift and in-frame SVs, respectively. (B) Deletions with frequencies > 0.1 mapped on the HBV genome are shown. Samples, where each deletion was detected with a frequency > 0.1, are described on the right of the deletion name. Red bars show the affected area of each deletion. The numbers at both ends of the bars signify the nucleotide positions of the breakpoints. (C) A scheme of breakpoint-specific fragments generated using Ion Proton sequencing. The gray bar at the top of the scheme represents the case-specific reference of case #06 with g.2449_488del (Δζ). The vertical bar at the middle of the scheme represents the sequence depth at each position. The vertical bars represent a position containing SNVs with *T* (red) and *A* (green). The horizontal bars at the bottom represent mapped reads with Δζ detected in sample 06-3. (D) Images of agarose gel electrophoresis of 2 deletion-specific amplicons derived from deletion positive (+) and negative (−) samples. Left: Δζ. Right: Δι. Abbreviation: RT indicates reverse transcriptase; SV, structural variation; SNV, single nucleotide variant; TP, terminal protein.

HBeAg-positive samples; however, the precore G1899A and core promoter A1764G mutations were not associated with HBeAg status (Figure 5E). In addition, the G1896A mutation allele frequency and the HBeAg titer showed a significant inverse correlation (Supplementary Figure 2B, http://links.lww.com/HC9/A103).

In contrast, genome deletion in the precore region was not detected in any of the samples, but the deletion in the *C* region was frequently found in anti-HBe-negative samples in the current study (Figure 5D). In case #04, a focal deletion in the *C* region was identified in most viral clones at time-points 1 and 2, whereas the number of these defective clones decreased at time-point 3

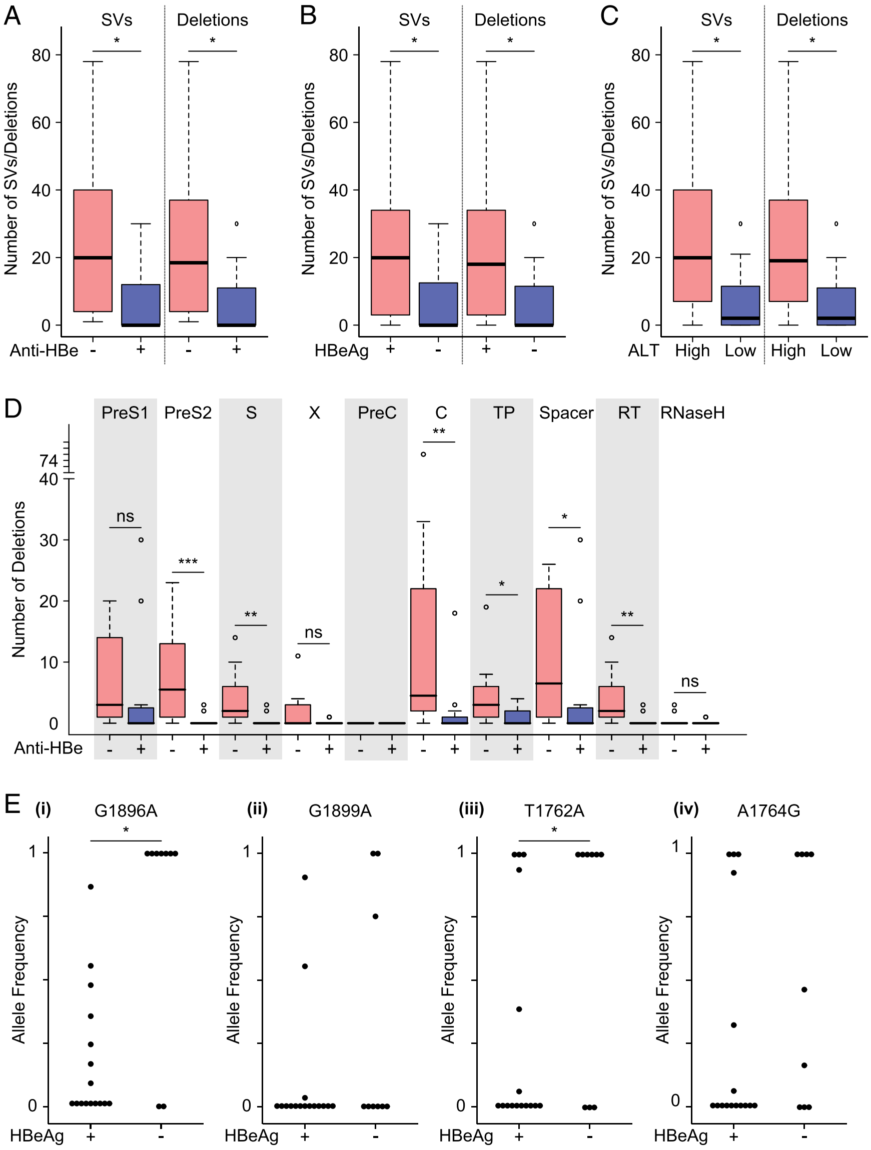

FIGURE 5 The association of genetic alterations in the HBV genome with clinical markers of chronic hepatitis B. Box plots showing the numbers of total SVs and deletions according to the presence of anti-HBe (A), HBeAg (B), and serum ALT level (C). (D) Association between anti-HBe status and the numbers of deletions per genetic region. Genome regions are listed at the bottom of the graph, such as preS1, *S*, and *X*. *X*-axis represents the clinical status of each marker. *Y*-axis represents the number of total SVs or deletions. The cutoff value for ALT was 40 IU/L. (E) Dot plots showing the allele frequencies of the G1896A (i), G1899A (ii), T1762A (iii), and A1764G (iv) mutations, according to the presence of HBeAg. * indicates p value <0.05, ** p value <0.01, *** p value <0.001, and "ns" no significant difference between the two groups. Abbreviation: ALT indicates alanine aminotransferase; SV, structural variation.

(Figure 6A). However, HBeAg was consistently positive in all samples.

These data suggest that viral clones with SVs/deletions are likely to seem under the condition of active hepatitis; however, focal genome deletion, which specifically affects seroconversion, was not identified.

Chronological dynamics of defective HBV clones

To clarify whether the defective HBV clones survived for a long time and whether they had any significance in the clinical course, we compared the entire genome sequences of all HBV clones detected at all time points.

In cases #04 and #06, in which active hepatitis was observed during the clinical course, we observed drastic changes in the frequencies of the defective clones. In case #04, defective clones with focal deletions in the *C* region occupied most of the population at time-points 1 and 2, while almost all defective clones were absent after the appearance of anti-HBe (time-point 3). Instead, various defective clones were detected at time-point 3 (Figure 6A). Case #06 had various defective clones throughout the clinical course, such as focal deletion of the *C* region at

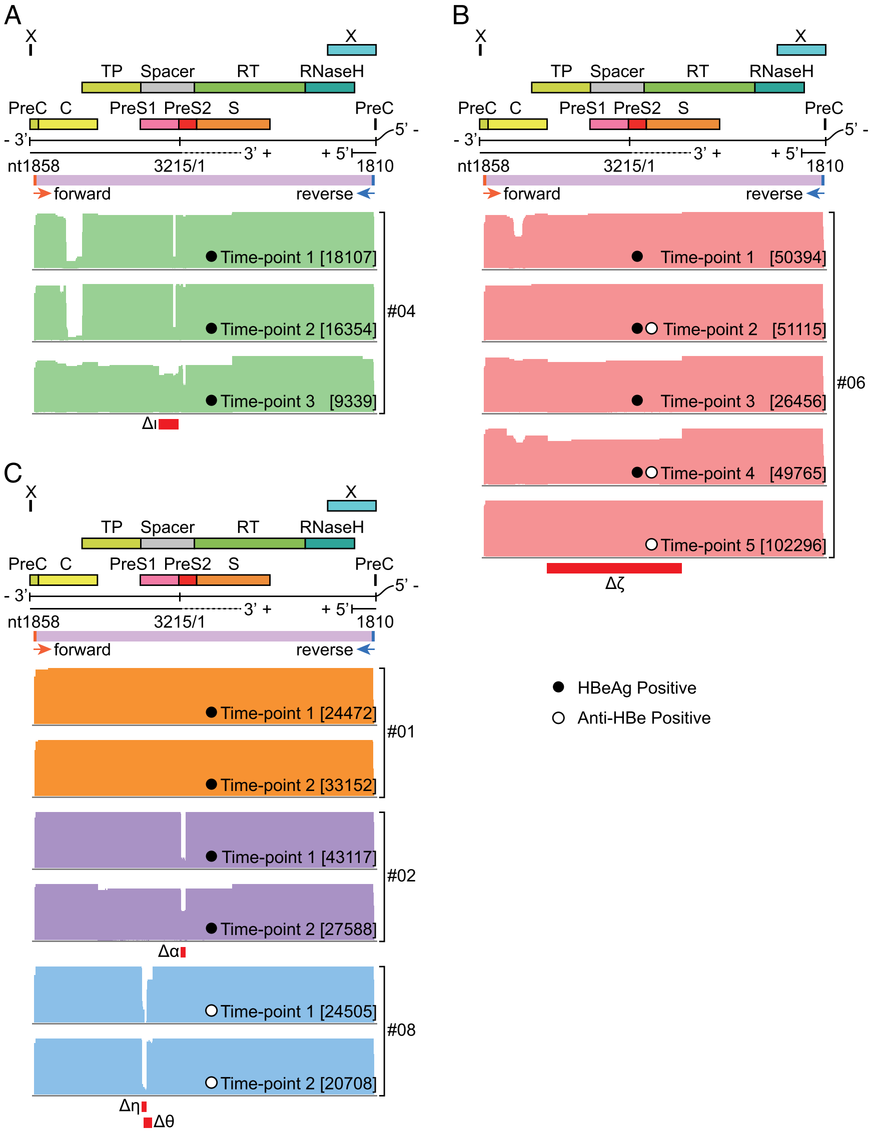

FIGURE 6 Comparison of the frequency of various defective HBV clones at the different time points. Schematic images of genetic regions of the reference sequence of the incomplete double-strand DNA of HBV are shown at the top of this panel. Under the HBV-DNA references, mapping images of each sample generated using NGMLR are given. In each case, the mapping of the first to last time points is arranged from top to bottom (2–5 time points per case). Black and white circles denote positive for HBeAg and anti-HBe, respectively, at each time-point. The numbers between brackets indicate the highest numbers of mapped reads among the nucleotide positions, including both full-length and defective clones. The red bars below the mapping images show the representative deletions detected in the sample. (A) Mapped images of SMRT sequencing at 3 time points in case #04. (B) Mapped images of 5 time points in case #06. (C) Mapped images of 6 samples from the representative cases, cases #01, #02, and #08. Abbreviation: RT indicates reverse transcriptase; SMRT indicates single-molecule real-time; TP, terminal protein.

time-point 1 and long deletions including the *C*, TP, preS1/2, and *S* regions at time-points 3 or 4 (Δζ). However, after seroconversion, these subpopulations became undetectable, suggesting that only full-length HBV clones could survive in the presence of the anti-HBe antibody in this case (Figure 6B). Meanwhile, Δζ was initially absent at time-point 1 in case #03 and emerged at time-point 2 after a >4-year-long observation (Supplementary Figure 3, http://links.lww.com/HC9/A104).

Figure 6C shows the 3 other representative cases. Case #01 possessed only full-length HBV clones at both time points, whereas the other cases had defective clones at all time points. Frequencies of viral populations with each genomic deletion altered with the different time-points in most cases. For example, in case #02, defective clones with short deletions in the preS2 region (g.11_52del; Δα) were the major population at time-point 1, whereas the frequency of these defective clones decreased at time-point 2. Interestingly, defective clones with long deletions over TP, spacer, preS1/2, *S*, and reverse transcriptase regions were detected at a low frequency at time-point 2. In contrast, in case #08, which had anti-HBe antibody throughout the observational period, defective clones with focal deletions in the preS1 region were continuously the major viral clones at time-points 1 and 2 (g.2861_2905del; Δη and g.2883_2958del; Δθ).

The sequence depths for the remaining 4 cases are shown in Supplementary Figure 3, (http://links.lww.com/HC9/A104). In case #07, 2 major defective clones with g.1763_1770del (Δγ) or g.1766_1771del (Δγ′) were detected at time-point 1, and none were detected at time-point 2. No major defective clones were detected throughout the observation period in cases #05, #09, and #10.

Altogether, the frequencies of defective HBV clones among all the HBV populations altered drastically throughout the observational period, according to the antiviral immunologic status in each case.

Phylogenetic landscape of quasispecies, with coexistence of full-length and defective HBV clones

To examine the genome-wide landscape of HBV quasispecies consisting of full-length and defective HBV clones, we performed phylogenetic analysis using the contiguous whole-genome sequence information of HBV clones in each serum sample. In this analysis, all SNVs in every viral clone were compared at single-viral clone resolution.

Figure 7A shows the phylogenetic tree generated using 200 viral clones, including 100 full-length and 100 defective g.11_52del (Δα) viruses at time-point 1 of case #02. This tree clearly shows that the coexisting full-length and defective clones were phylogenetically separated from each other. In case #04, where a double-deletion haplotype with g.2150_2305del (Δε) and g.3151_3177del (Δκ) was predominant, the distribution of each clone showed the same tendency (Figure 7B), suggesting that the full-length and defective clones accumulated new nucleotide alterations and clonally expanded independently. To examine the differences in the genetic diversity of the full-length clones, as compared with that of the defective clones, genetic distances were calculated among all clones in the population with full-length clones and compared to those with defective clones. The mean genetic distances, representing the degree of genetic diversity of each population, were longer in the full-length clones than in the defective clones, at time-point 1 in case #02, whereas the population with defective clones had a higher degree of diversity than the full-length clones at time-point 2 (Figure 7C). In case #04, genetic diversity did not significantly differ between the full-length and defective populations at any time point (Figure 7D). These data suggest that the diversity of viral clones is independent of SVs.

Evolution of defective HBV clones in the natural history

To evaluate the chronological dynamics of HBV infection, phylogenetic trees were constructed for each case. For the cases, in which the representative defective clones were absent or were of a very low frequency, most clones were separated from those at the other time points (time-points 2 or 3) in every case (Supplementary Figure 4, http://links.lww.com/HC9/A105).

Next, to clarify whether defective HBV clones could survive and spread independently of full-genome clones, we focused on the major defective HBV clones detected at multiple time points in each case and analyzed their chronological dynamics in years of natural clinical course. The phylogenetic tree described in Figure 8A was constructed of 100 randomly selected full-length and 100 Δα-defective clones at each time-point in case #02. In this tree, the viral cluster of full-length clones at time-point 2 overlaps with a part of that at time-point 1, while most Δα clones at time-point 2 are genetically close to those at time-point 1. Similarly, double-defective clones with both g.2150_2305del (Δε) and g.3151_3177del (Δκ) in case #04 accumulated mutations independently from full-length clones and finally formed independent viral clusters that were genetically distant from full-length HBV clones (Figure 8B). In case #08, g.2861_2905del (Δη) and g.2883_2958del (Δθ) defective clones were predominant, but no full-length clones were detected at the first time-point. Background nucleotide sequences of these defective clones are genetically different from each other, and these 2 populations have evolved independently at time points 1 and 2. Interestingly, full-length clones newly seemed at time-point 2, which formed a new cluster different from the preexisting 2 clusters of defective clones (Supplementary Figure 5A, http://links.lww.com/HC9/A106). In these cases, both full-length and defective clones are considered to evolve independently during the natural clinical course.

The viral dynamics in case #04 were quite complicated because of the presence of 4 different clusters of defective clones. Defective clones with g.24_56del (Δβ) alone and those with g.3020_3202del (Δι) were detected at time-points 1 and 2 at low frequencies. Notably, clones with double deletions (Δβ and Δι) emerged newly at time-point 3. To simplify their genetic relationship, we also constructed a phylogenetic tree that consisted only of clones with Δβ alone, Δι alone, and both together, at time-point 3 (Figure 8C). Δι clones were genetically distant from Δβ clones, and all the newly developed defective clones with double deletions overlapped with Δι clone cluster. In the absence of any evidence of viral recombination from the nucleotide sequence information, the new double-defective clones with Δβ and Δι are estimated to have evolved from preexisting Δι clones, which

京都大学
KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

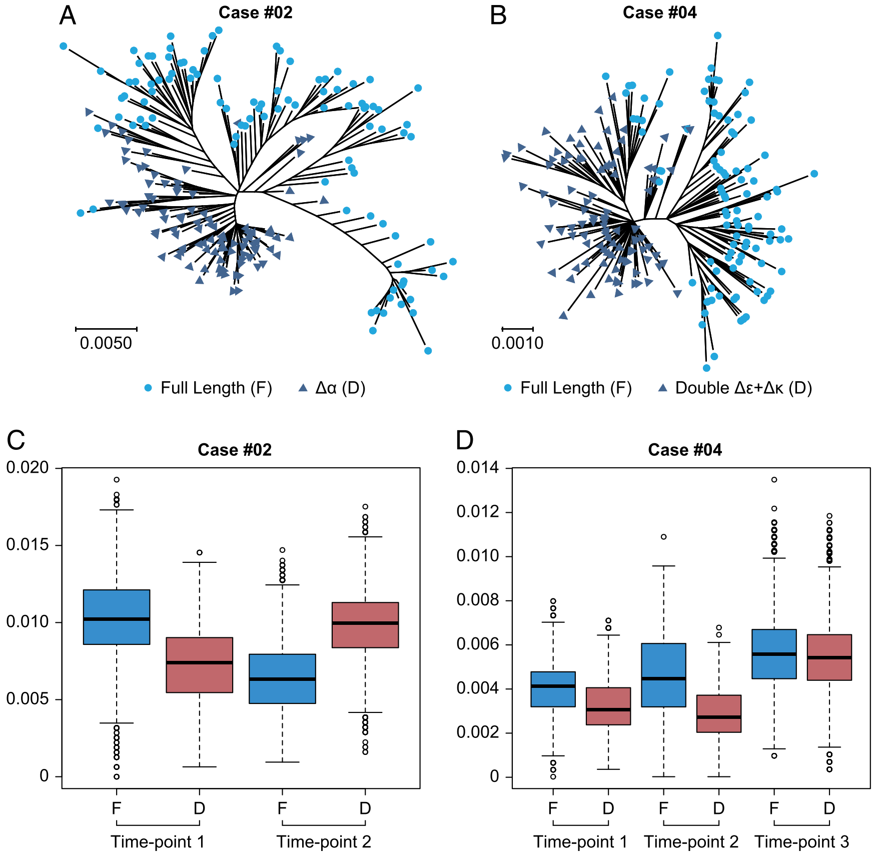

FIGURE 7 Quasispecies with coexistence of full-length and defective HBV clones. (A) Phylogenetic tree using 100 randomly selected clones each for full-length (blue circles) and defective (blue triangles) clones at time-point 1 of case #02. Defective clones with g.11_52del (Δα). (B) Phylogenetic tree of case #04 at time-point 1. Defective clones with g.2150_2305del and g.3151_3177del (Δε+Δκ). (C and D) Box plots showing the genetic distances among each cluster in case #02 (C) and case #04 (D). *X*-axis shows the cluster group. *Y*-axis shows the genetic distances. Abbreviation: F indicates full-length clones; D, defective clones.

have newly acquired the same deletions as Δβ, apart from preexisting defective clones with Δβ alone. We also constructed a phylogenetic tree of defective clones with Δε alone, Δκ alone, and double deletions (Δε and Δκ) at time-point 1. Based on the genetic distances, Δκ clones were estimated to have acquired Δε, apart from the preexisting Δε clones (Supplementary Figure 5B, http://links.lww.com/HC9/A106).

In contrast, some defective clones did not survive for long periods. In case #06, g.2449_488del (Δζ) was detected at a certain frequency at every time-point, except at the last time-point 5. The phylogenetic tree demonstrated that the cluster with full-length clones at time-point 1 overlapped with almost all defective clones with Δζ. A similar trend was observed at time points 3 and 4, whereas the major population evolved genetically (Figure 8D and Supplementary Figure 5C, http://links.lww.com/HC9/A106). These observations suggested that Δζ clones were derived from full-length clones at each time-point and gradually disappeared.

In some cases, defective clones, which were minor viral populations at the first time-point, became the major clones at the last time-point. In case #04, almost all viral populations consisted of defective clones at time-point 1, while the major clones were replaced at time-point 3 after the emergence of HBeAb. In case #08, the full-length clones were undetectable at first but emerged at the last time-point. In contrast, in case #07, g.1763_1770del (Δγ) clones and g.1766_1771del (Δγ′) clones were predominant at time-point 1, forming overlapped clusters, but became undetectable at time-point 2 (Supplementary Figure 5D, http://links.lww.com/HC9/A106). The major HBV sequence and major population constantly change over the years, even if no antiviral treatments are administered in the clinical course. The whole-genome nucleotide information of persistently infecting HBV changes every moment, not only at the SNV level but also at the level of SVs.

DISCUSSION

In the current study, we successfully determined contiguous whole-genome sequences of more than 100,000 HBV clones that simultaneously infected identical hosts. We found that genetically varied clones,

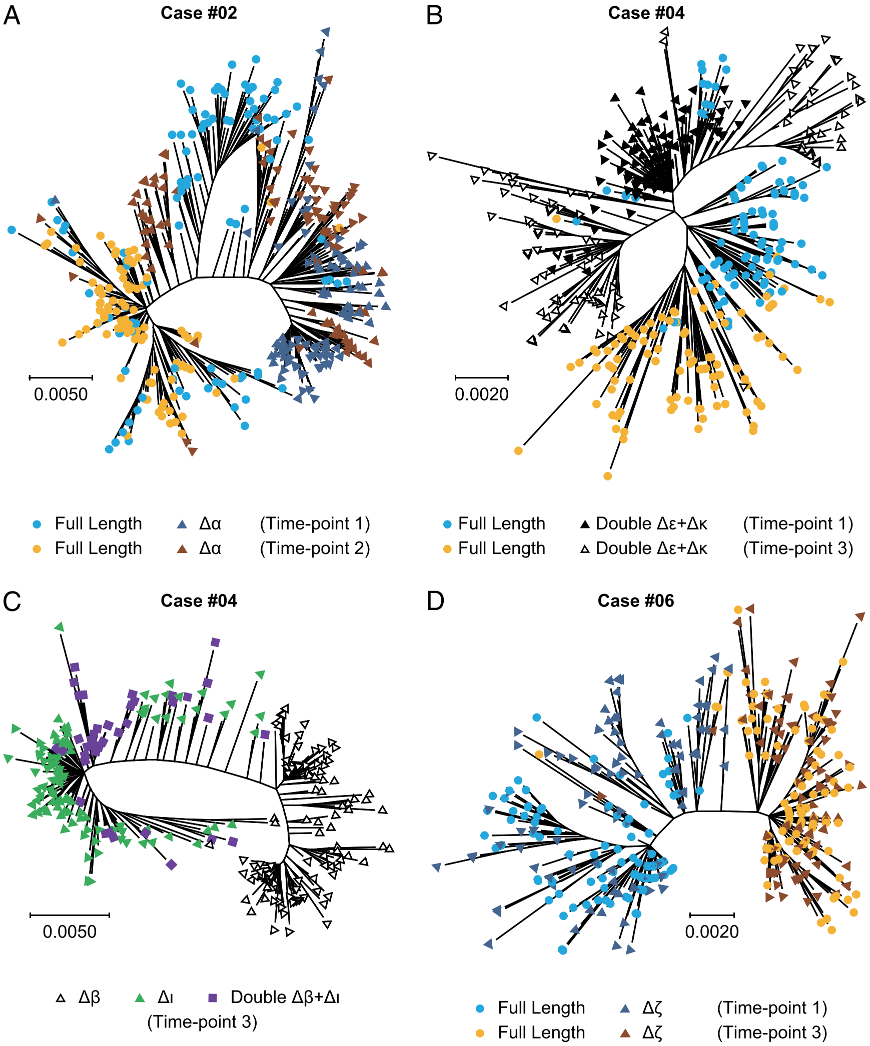

FIGURE 8 Phylogenetic analysis of the drastically changing HBV quasispecies during a persistent infection. Phylogenetic trees based on deletion-specific haplotyping at multiple time points (TPs) are shown. (A) Case #02. FL clones at TP1 (blue circles) and TP2 (orange circles) and defective clones with g.11_52del (Δα) at TP1 (indigo triangles) and TP2 (brown triangles) are shown. (B) Case #04. FL (TP1: blue circles and TP3: orange circles) and defective clones with both g.2150-2305del (Δε) and g.3151_3177del (Δκ) (TP1: black triangles and TP3: white triangles) are included in this tree. (C) Case #04. Just specific defective clones with g.24_56del (Δβ), clones with g.3020_3202del (Δι), and those with both Δβ and Δι at TP3 are focused on. Green triangles represent clones with only Δβ, purple squares with only Δι, and gray triangles with both Δβ and Δι. D. TP1 and TP3 of case #06. Blue circles, FL clones at TP1; orange circles, FL clones at TP3; indigo triangles, defective clones with g.2449_488del (Δζ) at TP1; and brown triangles, Δζ clones at TP3. Abbreviation: FL indicates full-length clones, TP, terminal protein.

including full-length and defective HBV clones, coexisted simultaneously in the serum of most patients. These varied HBV clones form heterogeneous quasispecies. To date, viral quasispecies has been mainly investigated using SNVs, including drug-resistant variants or vaccine-escape mutations.[27–35] In contrast, our current study demonstrates that HBV quasispecies is formed not only by SNVs but also by various SVs, mainly genome deletions. Deletions increase the diversity HBV genome, increasing the complexity of HBV quasispecies.

Although short deletions in HBV genome sequences have been reported earlier, these reports were based on conventional short-read next-generation sequencing data, which might have overlooked many structural alterations, including lengthy insertions and deletions.[15,36–42] In addition, conventional short-read sequencing cannot determine the continuous genome sequence of each HBV clone, which is one of the most fatal obstacles to elucidating the landscape of HBV quasispecies and its evolutionary process in natural history. Previous studies have shown the presence of lengthy deletions, so-called spliced variants, using conventional northern blot, PCR-based electrophoresis, and SMRT sequencing method.[10–13,16,43] However, the real dynamics of HBV quasispecies are unclear because a comprehensive analysis of the whole HBV genome in cases with chronic infection at multiple time points has not been performed. In the current study, by applying extremely long-read deep sequencing methodology with the PacBio platform in combination with multiple bioinformatics tools to HBV genome investigation, we achieved continuous whole-genome determination of HBV clones with various deletions (defective HBV clones), along with their clonal proportion in the total viral population, which is a breakthrough in investigating HBV quasispecies. Although DNA fragment integrated into the human genome was not excluded from the analysis, most of the DNA sequence spanning 3000 contiguous bases determined by SMRT sequencing was considered to be from relaxed circular DNA of viral particles in the serum.

Using this novel methodology, we first compared whole-genome HBV sequences of clinical cases with immunologically different statuses. In our analysis, sera positive for anti-HBe and negative for HBeAg, which are in the postseroconversion status, possessed a significantly lower frequency of SVs among the total HBV clones. As reported, HBV quasispecies is considered to be more diverse in HBV-infected cases with active hepatitis than in cases in the immunotolerant phase, in terms of both SNVs and short indels.[44–46] In cases of advanced liver disease, such as cirrhosis and hepatocellular carcinoma, the diversity is much higher than that in nonadvanced chronic hepatitis B.[47] Our study showed similar results to that observed in previous reports on HBV quasispecies containing all-sized SVs. In general, before achieving seroconversion, hepatitis activity is high, with elevated serum ALT levels. Our data suggest that a highly active immunologic response might result in a significantly higher frequency of HBV genome deletions.

In phylogenetic analysis, we found various defective HBV clones at most time points in many cases, the evolutionary patterns of which differed by case. Some defective HBV clones temporarily arose and then disappeared before the next time-point, and are thus considered to be “junk” fragments that cannot survive for a long time. Other defective HBV clones survived and evolved for years independently of full-length HBV clones, while drastically changing their populations according to the host-virus immunologic status. In contrast, a well-known long deletion named SP1 spliced variant, Δζ, has been detected for years in some cases,[14] but it seems ephemeral because Δζ clones were born from full-length clones at each time-point. Even long-surviving defective clones can disappear, and in some cases, almost all defective clones become undetectable after seroconversion.

Interestingly, defective clones can be predominant during the natural clinical course without intensive antiviral therapy and can survive for years. Most of these predominant long-surviving defective clones have in-frame deletions (27–183 nucleotides) in viral genetic regions such as *C*, preS1, and/or preS2, which encode structural proteins. Deletions in the preS1 or preS2 regions also affect the spacer region in the polymerase-coding open reading frame, whereas critical regions for viral replication such as TP, reverse transcriptase, and RNaseH are completely conserved. Although the reason for the selection of the specific defective clones at a certain time-point during the clinical course could not be clarified, these novel defective clones can serve as a novel “genotype” with a different genome size if they are transmitted to others and established as the major viral clones in the other host. This phenomenon suggests the natural occurrence of diverse HBV genotypes worldwide, some of which possess different-sized DNA genomes.

The current study had several limitations. First, the number of patients analyzed in this study was limited because HBV-positive patients not administered antiviral treatment are quite rare. However, using the novel SMRT sequencing technology, we could determine ~800,000 viral clones from the patient samples and analyze their dynamics in detail. Second, the PCR procedure used in the library generation for SMRT sequencing could result in noise for the metagenome data. The current standard library preparation requires a large amount of DNA; thus, we performed minimal cycles of PCR for the HBV-DNA samples. In addition, considering the random selection nature of CCS reads from enormous numbers of HBV clones circulating in the host serum, the noise caused by PCR could be regarded as minimal. Owing to the rapid progress in sequencing technologies, the limitations of PCR-associated noise can be overcome in the near future.

In conclusion, using novel long-read single-molecular deep sequencing technology in the present study, we discovered that HBV species are formed not only by SNVs but also by various SVs, especially defective variants. Even in clinical observations without antiviral therapy, various defective HBV clones were found to have been generated in natural history, especially under the condition of active hepatitis. Interestingly, we found that these defective viral clones evolved independently of the full-length HBV clones in many cases. Some of these naturally occurring defective clones may be involved in the development of novel genotypes of this pandemic virus. The novel analysis platform of long-read whole-genome sequencing of HBV quasispecies established in this study could serve as the next-generation standard to investigate the genetic landscape of HBV, and thus, help shed light on the ecology of HBV.

ACKNOWLEDGMENTS

The authors thank Drs Fumiyasu Nakamura, Tadashi Inuzuka, Tomonori Matsumoto, Tomoyuki Goto, Masayuki Ueno, Ken Kumagai, Masako Mishima, Mari Teramura, and Shigeharu Nakano at the Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, for interpreting the data and their helpful advice; Dr Akihiro Sekine at the Department of Omics-based Medicine, Chiba University, for helping with the PacBio data analysis; Drs Yukio Osaki, Norihiro

Nishijima, Toru Kimura, and Ryuichi Kita at the Department of Gastroenterology and Hepatology, Osaka Red Cross Hospital, and Drs Satoru Seo, Hideaki Okajima, Shinji Uemoto, Kojiro Taura, and Etsuro Hatano at the Division of Hepatobiliary and Pancreatic Surgery and Transplantation, Kyoto University Hospital, for their support.

FUNDING INFORMATION

This work was supported by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research, KAKENHI [grant numbers JP17H04158, JP19K08443, and JP20K22899].

CONFLICT OF INTEREST

All authors declare no conflicts of interests.

ORCID

Soichi Arasawa https://orcid.org/0000–0001–7856–6774
Haruhiko Takeda https://orcid.org/0000–0002–8954–9133
Atsushi Takai https://orcid.org/0000–0002–5233–4120
Eriko Iguchi https://orcid.org/0000–0002–7921–1104
Yuji Eso https://orcid.org/0000–0003–4426–1491
Takahiro Shimizu https://orcid.org/0000–0001–8392–2526
Taiki Yamashita https://orcid.org/0000–0002–5465–534X
Yoshihide Ueda https://orcid.org/0000–0003–3196–3494
Hiroyuki Marusawa https://orcid.org/0000–0002–4286–2712
Hiroshi Seno https://orcid.org/0000–0002–8509–8128

REFERENCES

1. Magnius L, Mason WS, Taylor J, Kann M, Glebe D, Dény P, et al. ICTV virus taxonomy profile: hepadnaviridae. J Gen Virol. 2020;101:571–2.
2. Nassal M. HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. Gut. 2015;64:1972–84.
3. Shi Y, Zheng M. Hepatitis B virus persistence and reactivation. BMJ. 2020;370:m2200.
4. Seeger C, Mason WS. Molecular biology of hepatitis B virus infection. Virology. 2015;479-480:672–86.
5. Sadiea RZ, Sultana S, Chaki BM, Islam T, Dash S, Akter S, et al. Phytomedicines to target hepatitis B virus DNA replication: current limitations and future approaches. Int J Mol Sci. 2022;23:1617.
6. Littlejohn M, Locarnini S, Yuen L. Origins and evolution of hepatitis B virus and hepatitis D virus. Cold Spring Harb Perspect Med. 2016;6:a021360.
7. Gilbert C, Feschotte C. Genomic fossils calibrate the long-term evolution of hepadnaviruses. PLoS Biol. 2010;8:e1000495.
8. Osiowy C, Giles E, Tanaka Y, Mizokami M, Minuk GY. Molecular evolution of hepatitis B virus over 25 years. J Virol. 2006;80:10307–4.
9. Suh A, Brosius J, Schmitz J, Kriegs JO. The genome of a Mesozoic paleovirus reveals the evolution of hepatitis B viruses. Nat Commun. 2013;4:1791.
10. Suzuki T, Masui N, Kajino K, Saito I, Miyamura T. Detection and mapping of spliced RNA from a human hepatoma cell line transfected with the hepatitis B virus genome. *Proc Natl Acad* Sci USA. 1989;86:8422–6.
11. Terré S, Petit M-A, Bréchot C. Defective hepatitis B virus particles are generated by packaging and reverse transcription of spliced viral RNAs in vivo. J Virol. 1991;65:5539–43.
12. Günther S, Sommer G, Iwanska A, Will H. Heterogeneity and common features of defective hepatitis B virus genomes derived from spliced pregenomic RNA. Virology. 1997;238:363–71.
13. Sommer G, van Bömmel F, Will H. Genotype-specific synthesis and secretion of spliced hepatitis B virus genomes in hepatoma cells. Virology. 2000;271:371–81.
14. Kremsdorf D, Lekbaby B, Bablon P, Sotty J, Augustin J, Schnuriger A, et al. Alternative splicing of viral transcripts: the dark side of HBV. Gut. 2021;70:2373–82.
15. 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.
16. Betz-Stablein BD, Töpfer A, Littlejohn M, Yuen L, Colledge D, Sozzi V, et al. Single-molecule sequencing reveals complex genome variation of hepatitis B virus during 15 years of chronic infection following liver transplantation. J Virol. 2016;90:7171–83.
17. Yamashita T, Takeda H, Takai A, Arasawa S, Nakamura F, Mashimo Y, et al. Single-molecular real-time deep sequencing reveals the dynamics of multi-drug resistant haplotypes and structural variations in the hepatitis c virus genome. Sci Rep. 2020;10:2651.
18. Takeda H, Ueda Y, Inuzuka T, Yamashita Y, Osaki Y, Nasu A, et al. Evolution of multi-drug resistant HCV clones from pre-existing resistant-associated variants during direct-acting antiviral therapy determined by third-generation sequencing. Sci Rep. 2017;7:45605.
19. Lee GH, Wasser S, Lim SG. Hepatitis B pregenomic RNA splicing—the products, the regulatory mechanisms and its biological significance. Virus Res. 2008;136:1–7.
20. Abraham TM, Lewellyn EB, Haines KM, Loeb DD. Characterization of the contribution of spliced RNAs of hepatitis B virus to DNA synthesis in transfected cultures of Huh7 and HepG2 cells. Virology. 2008;379:30–7.
21. Huang C-C, Kuo T-M, Yeh C-T, Hu C-p, Chen Y-L, Tsai Y-L, et al. One single nucleotide difference alters the differential expression of spliced RNAs between HBV genotypes A and D. Virus Res. 2013;174:18–26.
22. Chen J, Wu M, Wang F, Zhang W, Wang W, Zhang X, et al. Hepatitis B virus spliced variants are associated with an impaired response to interferon therapy. Sci Rep. 2015;5:16459.
23. Lam AM, Ren S, Espiritu C, Kelly M, Lau V, Zheng L, et al. Hepatitis B virus capsid assembly modulators, but not nucleoside analogs, inhibit the production of extracellular pregenomic RNA and spliced RNA variants. Antimicrob Agents Chemother. 2017;61:e00680–00617.
24. Lim CS, Sozzi V, Luciani F, Littlejohn M, Revill PA, Brown CM. Quantitative analysis of the splice variants expressed by the major hepatitis B virus genotypes. Microb Genom. 2021;7:mgen000492.
25. Luk KC, Gersch J, Harris BJ, Holzmayer V, Mbanya D, Sauleda S, et al. More DNA and RNA of HBV SP1 splice variants are detected in genotypes B and C at low viral replication. Sci Rep. 2021;11:23838.
26. Kandpal M, Samal J, Biswas B, Negi A, Mishra VC, Tyagi N, et al. Enhanced hepatitis B virus (HBV) pre-genomic RNA levels and higher transcription efficiency of defective HBV genomes. J Gen Virol. 2015;96:3109–17.
27. Nishijima N, Marusawa H, Ueda Y, Takahashi K, Nasu A, Osaki Y, et al. Dynamics of hepatitis B virus quasispecies in

association with nucleos(t)ide analogue treatment determined by ultra-deep sequencing. PLoS ONE. 2012;7:e35052.
28. Hayer J, Rodriguez C, Germanidis G, Deléage G, Zoulim F, Pawlotsky J-M, et al. Ultradeep pyrosequencing and molecular modeling identify key structural features of hepatitis B virus RNase H, a putative target for antiviral intervention. J Virol. 2014;88:574–82.
29. Lowe CF, Merrick L, Harrigan PR, Mazzulli T, Sherlock CH, Ritchie G. Implementation of next-generation sequencing for hepatitis B Virus resistance testing and genotyping in a clinical microbiology laboratory. J Clin Microbiol. 2016;54:127–33.
30. Gencay M, Vermeulen M, Neofytos D, Westergaard G, Pabinger S, Kriegner A, et al. Substantial variation in the hepatitis B surface antigen (HBsAg) in hepatitis B virus (HBV)-positive patients from South Africa: Reliable detection of HBV by the Elecsys HBsAg II assay. J Clin Virol. 2018;101:38–43.
31. Chevaliez S, Rodriguez C, Poiteau L, Soulier A, Donati F, Darty-Mercier M, et al. Primary resistance of hepatitis B virus to nucleoside and nucleotide analogues. J Viral Hepat. 2019;26: 278–86.
32. Hayashi S, Murakami S, Omagari K, Matsui T, Iio E, Isogawa M, et al. Characterization of novel entecavir resistance mutations. J Hepatol. 2015;63:546–3.
33. Han Y, Zhang Y, Mei Y, Wang Y, Liu T, Guan Y, et al. Analysis of hepatitis B virus genotyping and drug resistance gene mutations based on massively parallel sequencing. J Virol Methods. 2013; 193:341–7.
34. Rajoriya N, Combet C, Zoulim F, Janssen HLA. How viral genetic variants and genotypes influence disease and treatment outcome of chronic hepatitis B. Time for an individualised approach? J Hepatol. 2017;67:1281–97.
35. Bruni R, Villano U, Taffon S, Equestre M, Madonna E, Chionne P, et al. Retrospective analysis of acute HBV infections occurred in 1978–79 and 1994– 95 in North-East Italy: increasing prevalence of BCP/pre-core mutants in sub-genotype D3. BMC Infect Dis. 2020;20:78.
36. Wu I-C, Liu W-C, Chang T-T. Applications of next-generation sequencing analysis for the detection of hepatocellular carcinoma-associated hepatitis B virus mutations. J Biomed Sci. 2018;25:51.
37. Teng C-F, Huang H-Y, Li T-C, Shyu W-C, Wu H-C, Lin C-Y, et al. A next-generation sequencingbased platform for quantitative detection of hepatitis B virus pre-S mutants in plasma of hepatocellular carcinoma patients. Sci Rep. 2018;8:14816.
38. Anastasiou OE, Widera M, Verheyen J, Korth J, Gerken G, Helfritz FA, et al. Clinical course and core variability in HBV infected patients without detectable anti-HBc antibodies. J Clin Virol. 2017;93:46–52.
39. Anastasiou OE, Widera M, Westhaus S, Timmer L, Korth J, Gerken G, et al. Clinical outcome and viral genome variability of hepatitis b virus–induced acute liver failure. Hepatology. 2019; 69:993–1003.
40. Cheng J-H, Liu W-C, Chang T-T, Hsieh S-Y, Tseng VS. Detecting exact breakpoints of deletions with diversity in hepatitis B viral genomic DNA from next-generation sequencing data. Methods. 2017;129:24–32.
41. Jia J, Liang X, Chen S, Wang H, Li H, Fang M, et al. Next-generation sequencing revealed divergence in deletions of the preS region in the HBV genome between different HBV-related liver diseases. J Gen Virol. 2017;98:2748–58.
42. 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.
43. Sauvage V, Boizeau L, Candotti D, Vandenbogaert M, Servant-Delmas A, Caro Vr, et al. Early MinION™nanopore single-molecule sequencing technology enables the characterization of hepatitis B virus genetic complexity in clinical samples. PLOS ONE. 2018;13:e0194366.
44. Yang Z-T, Huang S-Y, Chen L, Liu F, Cai X-H, Guo Y-F, et al. Characterization of full-length genomes of hepatitis B virus quasispecies in sera of patients at different phases of infection. J Clin Microbiol. 2015;53:2201–14.
45. Gao S, Duan Z-P, Coffin CS. Clinical relevance of hepatitis B virus variants. World J Hepatol. 2015;7:1086–96.
46. Xue Y, Wang MJ, Huang SY, Yang ZT, Yu DM, Han Y, et al. Characteristics of CpG Islands and their quasispecies of full-length hepatitis B virus genomes from patients at different phases of infection. SpringerPlus. 2016;5:1630.
47. Yamani LN, Yano Y, Utsumi T, Juniastuti, Wandono H, Widjanarko D, et al. Ultradeep sequencing for detection of quasispecies variants in the major hydrophilic region of hepatitis B virus in Indonesian patients. J Clin Microbiol. 2015;53:3165–75.

How to cite this article: Arasawa S, Takeda H, Takai A, Iguchi E, Eso Y, Shimizu T, et al. Evolutional transition of HBV genome during the persistent infection determined by single-molecule real-time sequencing. Hepatol Commun. 2023;7:e0047. https://doi.org/10.1097/HC9.0000000000000047