Observational Study

Medicine®

OPEN

Integrating nested PCR with high-throughput sequencing to characterize mutations of HBV genome in low viral load samples

Xianjun Wang, MSc[a,*], Lihui Xu, MSc[a], Yueming Chen, MSc[a], Anbing Liu, MSc[a], Liqian Wang, MSc[b], Peisong Xu, BS[c], Yunhui Liu, MSc[c], Lei Li, MSc[c], Fei Meng, MSc[c]

Abstract

Due to the low viral load of hepatitis B virus (HBV) in plasma samples, conventional techniques have limitations to the detection of antiviral resistance mutations. To solve the problem, we developed a fast, highly sensitive, and accurate method to sequence the HBV whole-genome sequencing in plasma samples which had various viral loads from very low to high.

Twenty-one plasma samples were collected from patients who were carriers of HBV from the Hangzhou First People's Hospital. Two pairs of conserved, overlapping, nested primers were used to amplify and sequence the whole HBV genome in 8 plasma samples with different viral loads. High-throughput sequencing was performed on Illumina MiSeq platform. Concomitantly, 3 samples were directly sequenced without PCR amplification. We compared amplicon-sequencing with direct sequencing to develop a method for amplifying and characterizing the whole genome of HBV.

HBV genome was amplified from all samples and verified by Sanger sequencing, regardless of the viral loads. Sequencing results revealed that only a few reads were mapped to the HBV genome following direct sequencing, while the amplicon-sequencing reads had a good coverage and depth. We identified 50 intrahost single nucleotide variations (iSNVs), 14 of which were low frequency mutations. Interestingly, iSNVs were more common in low viral load samples than in high viral load samples, and mutations in the reverse transcriptase (RT) region were most prevalent.

We conclude that amplicon-sequencing is not only a practical method to detect HBV infection with a high sensitivity and accuracy but also enables to detect mutations in the HBV genome in low viral load samples from HBV-infected patients. Thus, our findings provide a new diagnosis method of HBV infection, which is capable of detection of low frequent mutations in low viral load samples.

Abbreviations: HBV = hepatitis B virus, HTS = high-throughput sequencing, iSNVs = intrahost single nucleotide variations, NAs = nucleoside analogues, RT = reverse transcriptase, SNVs = single nucleotide variations.

Keywords: amplicon-sequencing, HBV infection, high-throughput sequencing, iSNV, low frequent mutations, low viral load sample

1. Introduction

Hepatitis B virus (HBV) infection is a severe global health problem, and there is a high prevalence of HBV infection in China with approximately 93 million HBV carriers, which results in approximately 330,000 deaths annually.[1] Early and timely diagnosis of HBV infection is of great importance in the treatment of HBV infections.

HBV infection consists of viral attachment, entry and release of the nucleocapsid into the cytoplasm, nucleus import of the viral genome, followed by genome replication in the nucleus.[2] Progression of liver diseases caused by HBV infection is fostered by HBV DNA replication.[3] HBV genome is a partially double-stranded DNA, consisting of 4 overlapping open reading frames encoding viral polymerase (P), surface antigen (S), core (C), and X protein.[4] Due to the lack of polymerase proofreading activity of the viral polymerase and a high replication rate, HBV genome replication has a higher mutation rate than other DNA viruses. Many studies have reported that mutations in the HBV genome are associated with drug resistance, liver disease progression, and immune escape.[5,6] As a result, sequencing of the whole HBV genome is critical for understanding the correlation between mutations and liver disease progression.

To date, several conventional techniques have been used to sequence viral genomes and detect the mutations, such as DNA microarray, Sanger sequencing, and qPCR. However, none of these can fully satisfy the requirements for genetic diagnosis in terms of their specificity, sensitivity, or accuracy. This is in particular important when these methods are applied to low viral load samples and to detection of low frequent mutations. The mutation frequencies of Sanger sequencing and qPCR are >20% and >10% to 15%, respectively.[7]

Editor: Durga Tripathi.

Funding: This work received financial support from the Natural Science Foundation of Zhejiang Province (no. LY13H190003).

Authors' contributions: XW conceived and designed the experiments. LX and YC collected samples from the experiments. AL and LW performed the experiments. PX and FM analyzed the data. YL and LL performed sequencing. XW wrote the paper. FM checked and revised the manuscript.

The authors have no conflicts of interest to disclose.

[a] Clinical Laboratory, Hangzhou First People's Hospital, [b] Zhejiang Chinese Medical University, [c] Department of Research Service, Zhiyuan Inspection Medical Institute, Hangzhou, Zhejiang, People's Republic of China.

[*] Correspondence: Xianjun Wang, Clinical Laboratory, Hangzhou First People's Hospital, No. 261, Huansha Road, Hangzhou, Zhejiang, People's Republic of China (e-mail: wangxj0525@126.com).

Medicine (2017) 96:30(e7588)

Received: 18 October 2016 / Received in final form: 20 June 2017 / Accepted: 30 June 2017

http://dx.doi.org/10.1097/MD.0000000000007588

Table 1

The results of clinical information and quantification of HBV DNA or viral titer.

No.	Age	Clinical diagnose	Clinical symptoms	Entecavir treatment	Total bilirubin (μmol/L)	Child–Turcotte–Pugh	Qubit* (ng/μL)	q-PCR† (copies/mL)
H2	23	Chronic active hepatitis B	Debilitation	6 months	5.2	5	0.946	2.76×10^3
H1	32	Chronic active hepatitis B	Debilitation	15 months	12.5	5	4.08	3.157×10^3
H10	41	Chronic active hepatitis B	Jaundice	10 months	45.8	6	13.2	5.32×10^3
A2	26	Chronic active hepatitis B	Epigastric discomfort	6 months	7.6	5	3.9	6.43×10^3
H15	28	Chronic active hepatitis B	Debilitation	9 months	6.3	5	3.66	7.93×10^3
H8	37	Hepatitis B surface antigen carriers	Epigastric discomfort	Untreated	7.2	5	3.72	8.3×10^3
H19	52	Chronic active hepatitis B	Debilitation	8 months	6.8	5	1.93	9.69×10^3
H3	27	Hepatic fibrosis	Debilitation, jaundice	3 months	43.6	7	1.3	1.19×10^4
H5	47	Chronic active hepatitis B	Mild ascites	11 months	23.6	8	0.52	1.28×10^4
H20	27	Chronic active hepatitis B	Debilitation	6 months	8.3	5	2.34	1.72×10^4
H6	41	Chronic active hepatitis B	Moderate ascites	12 months	46.5	8	1.74	1.95×10^4
A1	24	Chronic active hepatitis B	Debilitation	2 months	5.1	5	1.25	2.52×10^4
H18	27	Hepatic fibrosis	Epigastric discomfort	1 months	5.7	6	2.94	2.67×10^4
H22	40	Chronic active hepatitis B	Mild ascites	6 months	36.8	8	0.598	3.02×10^4
H17	26	Hepatic fibrosis	Mild ascites	1 months	35.9	6	2.4	3.5×10^4
H7	28	Chronic active hepatitis B	Debilitation	3 months	19.6	6	1.47	5.48×10^4
A5	29	Chronic active hepatitis B	Jaundice	5 months	58.3	6	1.63	1.68×10^5
H21	26	Chronic active hepatitis B	Debilitation	5 months	22.5	6	0.72	4.81×10^5
H16	29	Chronic active hepatitis B	Jaundice	9 months	47.1	6	4.6	2.42×10^6
H11	31	Chronic active hepatitis B	Jaundice	5 months	52.7	6	0.366	1.61×10^7
H14	2	Chronic active hepatitis B	Debilitation	6 months	9.3	5	1.28	4.81×10^7

ELISA = enzyme linked immunosorbent assay.

* Qubit indicates that the concentration of whole genome (including host and virus) which was extracted by PureLink Viral RNA/DNA Mini Kit.

† q-PCR indicates that the concentration of HBV genome which was measured by ABI 7500.

In recent years, high-throughput sequencing (HTS) has proven to be a powerful tool in diagnosis of virus infections.

HTS has been used to sequence the complete HBV genome for identification of virulence factors, drug resistance genes and characterization of single nucleotide variations (SNVs) of the HBV genome.[8,9] Additionally, previous studies have demonstrated that HTS is not sensitive enough to detect low levels of viral genome presented in individual plasma samples.[10] Increasing the copy numbers of the viral genome using PCR is thus a key step in sample preparation. However, it is very difficult to amplify the complete HBV genome, because the HBV genome involves a partially double-stranded DNA molecule with 2 breaks. In 1995, Gunther et al[11] proposed a 1-step PCR method with a single pair of primers at the termini of the breaks to amplify the HBV genome; however, it is not sensitive enough to detect low viral load of HBV in plasma samples. Subsequently, a highly sensitive amplification method, nested PCR, was proposed in a follow up study, especially for low viral load samples.[12–14]

In this study, we successfully integrated nested PCR using 2 pairs of conserved, overlapping nested primers with HTS to develop a fast, highly sensitive, and accurate method for whole-genome sequencing of HBV in plasma samples with various viral loads.

2. Materials and methods

2.1. Patients and samples

We collected 21 plasma samples from patients who were HBV carriers in the First People's Hospital of Hangzhou from March 2015 to December 2015. The hospital ethics committee agreed to conduct the study and the ethical number was No. 051-01. All samples were conducted with conventional detection of HBV infection during preparation of plasma, and the concentration of HBV was initially determined using ELISA (Table 1).

2.2. DNA extraction and quantification

HBV DNA was extracted from 200 μL of plasma sample using the PureLink Viral RNA/DNA Mini Kit (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions (https://tools.thermofisher.com/content/sfs/manuals/purelink_viral_rna_dna_man.pdf) with modification of using 25 μL of sterile water to elute DNA. DNA concentration was quantified using the Qubit dsDNA HS Assay Kit (Invitrogen) (https://tools.thermofisher.com/content/sfs/manuals/Qubit_dsDNA_HS_Assay_UG.pdf). HBV viral load was determined by the Diagnostic Kit for Quantification of Hepatitis B Virus DNA (Puruikang, Beijing, China) using ABI7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, CA), according to the manufacturer's instructions.

2.3. Primers and amplification

The complete HBV genome was amplified using nested primers (Table 2) and validated with Sanger sequencing using specific primers.[11,15] The HBV genome was divided into 2 fragments (Fig. 1) and both rounds of nested PCR were performed in 25 μL reaction system. The reaction contained 2 μL of template DNA, 0.5 μL of each 10 μM primer, 2.5 μL of 10× LA Taq Buffer ($MgCl_2$-free), 4 μL of dNTP mixture (2.5 mM each), 2.5 μL of 25 mM $MgCl_2$, and 0.25 μL of TaKaRa LA Taq (5 units/μL). Thermal cycle parameters included predenaturing at 94°C for 5 minutes, followed by 35 cycles (first round of PCR) or 20 cycles (second round of PCR) of denaturation at 94°C for 30 seconds, annealing at 58°C for 30 seconds, extension at 72°C for 3 minutes, and finally extension for 10 minutes at 72°C. PCR products were analyzed in 1.5% agarose gel electrophoresis and cleaned up with Qiagen PCR Purification Kit (Qiagen, Valencia, CA). To validate the sequences of HBV genome, Sanger sequencing was performed with the ABI 3730XL DNA Analyzer

Table 2

The HBV-specific primers were used for amplifying and sequencing in 2 fragments.

Frag*	Primer	Sequences (5′–3′)	Position (nt)	Size† (bp)
A	P3	CTCGCTCGCCCAAATTTTTCACCTCTGCCTAATCA	1825–1841	2092
	AR1	ACAGTGGGGGAAAGC	759–745	
	AR2	AGAAACGGRCTGAGGC	702–687	
B	P4	CTGGTTCGGCCCAAAAAGTTGCATGGTGCTGG	1823–1806	1320
	AF1	GTCTGCGGCGTTTTATC	419–435	
	AF2	TGCCCGTTTGTCCTCTA	503–519	

* Whole genome of HBV was divided into 2 fragments: A and B. For the fragment A, P3 and AR1 were mixed in the first round of PCR, P3 and AR2 were mixed in the second round of PCR. For the fragment B, P4 and AF1 were mixed in the first round of PCR, P4 and AF2 were mixed in the second round of PCR.

† Size represented the products of second round of nested PCR.

(Applied Biosystems) using BigDye Terminators V3.1 according to the manufacturer's instructions.

2.4. Library construction and HTS

Nested PCR products were purified, and the concentrations were determined using the Qubit dsDNA HS Assay Kit (Invitrogen). Two nested PCR fragments from the same sample were mixed with equimolar quantities. Mixed PCR products were used to construct libraries using the Nextera XT DNA Sample Prep Kit (Illumina, San Diego, CA), according to the manufacturer's instructions. The quality of the libraries was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) with a high-sensitivity DNA chip. Libraries were subjected to HTS on the Illumina MiSeq (Illumina) platform to generate 2 × 250-bp paired-end reads.

To evaluate the efficiency of amplicon-sequencing, 3 different viral loads (high, mediate, and low) of HBV plasma samples were directly sequenced without PCR[14,16] on the Illumina MiSeq platform. In order to eliminate biases in sequencing to the maximum, we also used Nextera XT DNA Sample Prep Kit to construct libraries for direct sequencing.

2.5. Data analysis

FASTQ-format raw reads were generated on the Illumina MiSeq sequencer, and a quality control was processed simultaneously.

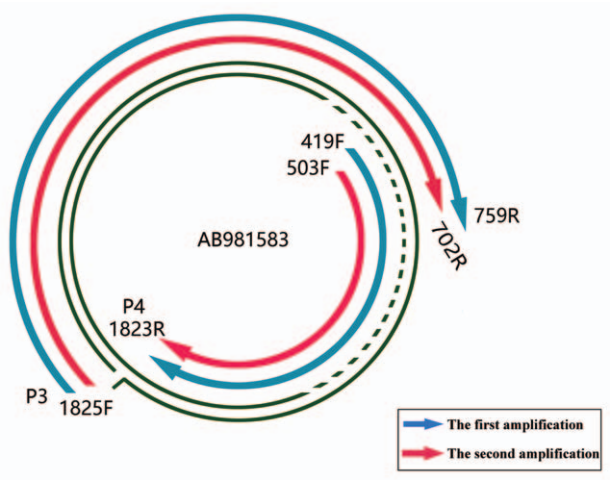

Figure 1. The structure of HBV genome and nested PCR. The blue arrow indicates the first round of amplification and the red arrow indicates the second round of amplification. The primers are depicted along with the HBV genome.

To reduce the false discovery rate of mutations, we trimmed the first 10 bases and the last 30 bases from every >Q20 read. The remaining properly paired reads were mapped against the HBV B-type reference genome (AB981583) using Bowtie2.[17] Intra-host single nucleotide variations (iSNVs) were called using the default parameters in the mapping algorithm.

3. Results

3.1. HBV detection and selection of sequencing samples

Primary HBV concentrations of the 21 plasma samples were determined by ELISA. Subsequently, viral genomic DNA was extracted, and the genome copy number was quantified using Qubit and qPCR. The concentrations and viral loads (genome copies) ranged from 0.366 ng/μL to 4.6 ng/μL and 10^3 to 10^7 (Table 1), respectively.

Given the bias of HTS, we selected 8 samples of various viral loads, according to the results of ELISA and qPCR. Three low viral load samples (H1, H2, H10), 3 moderate viral load samples (H20, H21, H22), and 2 high viral load samples (H11, H14) were selected.

3.2. Nested PCR and Sanger sequencing

LA Taq (Takara, Kyoto, Japan) and Phusion High-Fidelity PCR Master Mix with HF Buffer (New England BioLabs, Allschwil, Switzerland) were used to amplify HBV genome. The whole HBV genome was successfully amplified with 2 pairs of nested primers using LA Taq in the aforementioned 8 plasma samples, while the efficiency of the Phusion polymerase was lower than that of the LA Taq (Fig. 2). After 2 rounds PCR products were amplified at 2092 and 1320 bp, respectively. The results of the Sanger sequencing confirmed that the sequences of the PCR products belonged to HBV. Thus, these results suggest that the nested PCR is an appropriate method to obtain HBV whole-genome sequence, regardless of the initial HBV loads.

3.3. Overview of HBV amplicon-sequencing

HTS resulted in an average of 5092 clean reads (range from 2362 to 6792, Table 3) per sample after the low quality reads were removed. The clean paired-end reads were aligned to the reference sequence of the HBV genome, and the alignment rate approached 80% except for samples H1 and H10. The degree of coverage and depth for each sample were high and similar in magnitude (Fig. 3). However, positions 1823 to 1841 had an insufficient depth, which was caused by the gap in the HBV genome structure. Additionally, positions of 500 to 750 also had a lower depth because this DNA segment represents the overlap

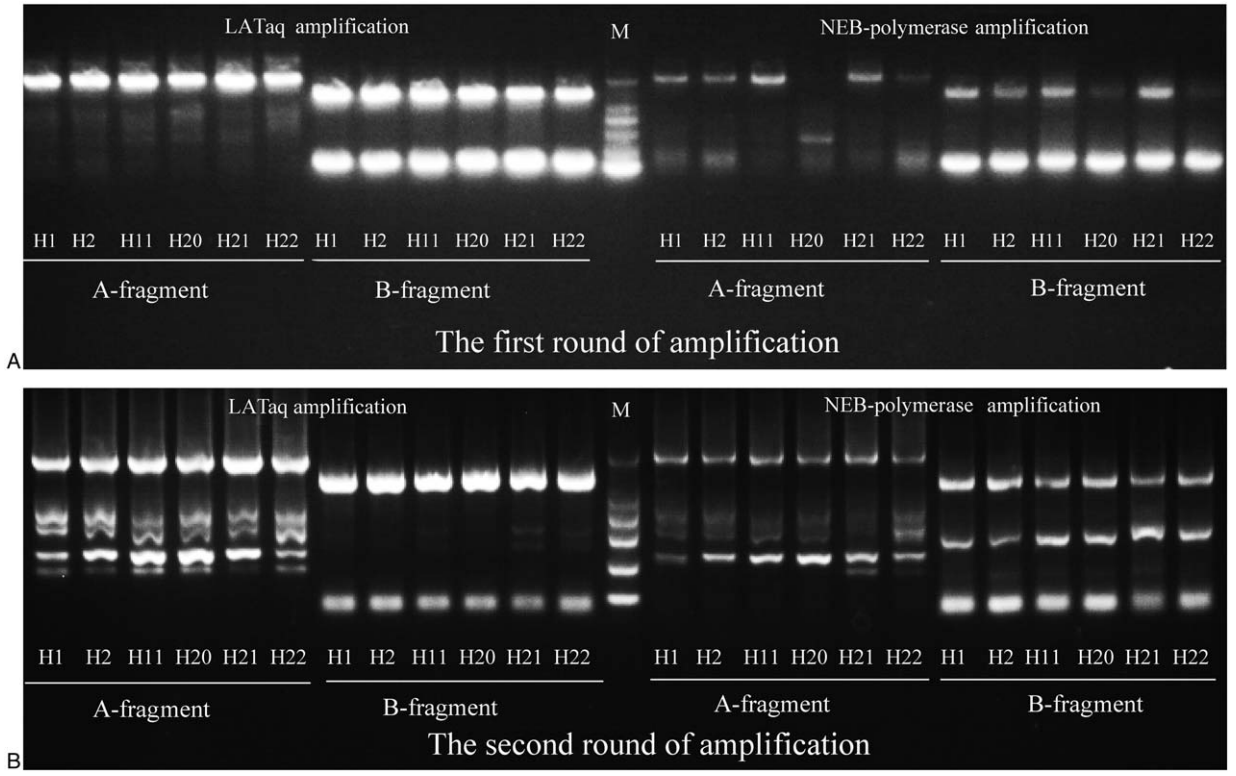

Figure 2. The results of HBV genomic amplification. (A) The results of first amplification. (B) The results of second amplification. M: marker (2000 bp).

between adjacent amplicons, but this was determined to be inevitable and acceptable. Therefore, the nested PCR products obtained from different viral load samples were suitable for HTS based on account of its uniformity and unbias.

3.4. HTS without PCR

Sample H1 (low viral load sample), H20 (mediate viral load sample), and H11 (high viral load sample) were directly sequenced without amplification using nested PCR. After trimming the low quality reads, a mean of 1.77 million clean reads (H1: 1.6 M reads, H20: 2.0 M reads, H11: 1.7 M reads) were used for further bioinformatics analysis. Human genomes were wiped out from clean reads and the rest of reads were mapped to the database of the viral genome. Unexpectedly, the number of reads mapped to HBV genome was very low, with only 10 and 4 reads mapped in the high viral load sample H11 and the mediate viral load sample H20, respectively. Furthermore, for the low viral load sample H1, there were no reads that could be mapped to the HBV genome.

Table 3

The results of amplicon-sequencing for each sample.

Sample	Clean reads	Alignment rate (%)	Depth	iSNV no.
H1	3400	37.53	46.9	8
H2	5882	79.07	66.2	9
H10	5948	38.28	44.6	13
H20	4302	82.19	55.7	0
H21	6792	81.35	50.2	14
H22	6206	85.05	69.5	2
H11	2362	79.34	36.7	4
H14	5840	85.03	70.2	0

iSNV = intrahost single nucleotide variation.

3.5. Analysis of iSNVs

Based on the results of the alignment with Bowtie2, iSNVs were identified with SAMTOOLS[18] (version 1.1). As shown in Table 3, 6 of the 8 samples contained iSNVs, with the number per sample ranging from 2 to 14. The sample H14 and H20 were failed to identify iSNVs. The average depth of mutation was 44.7 (range 18–92) (Table 4). To further characterize the HBV mutations, we annotated a total of 50 iSNVs. The lowest frequency of mutation was 0.02 and the highest frequency was 0.75. In some cases, the reference sequence mutated to 2 or more different bases, though this occurred in a relatively low frequency.

4. Discussion

Although strategies used to detect HBV infection have been improved with new approaches such as direct sequencing and qPCR,[19–22] it remains a critical issue to detect HBV genome in serum or plasma samples that contain low viral loads. To this end, we developed a method, integrating 2-stage nested PCR with HTS to overcome this difficulty. HBV genome copies were determined using qPCR and nested PCR products were extracted from agarose gel of electrophoresis and subjected to Sanger sequencing.

In this study, to sequence the whole genome of HBV in samples with various virus loads, we designed 2 strategies, including 1-step PCR using a single pair of primers[11] and nested PCR using 2 pairs of nested primers. Unfortunately, using a single pair of

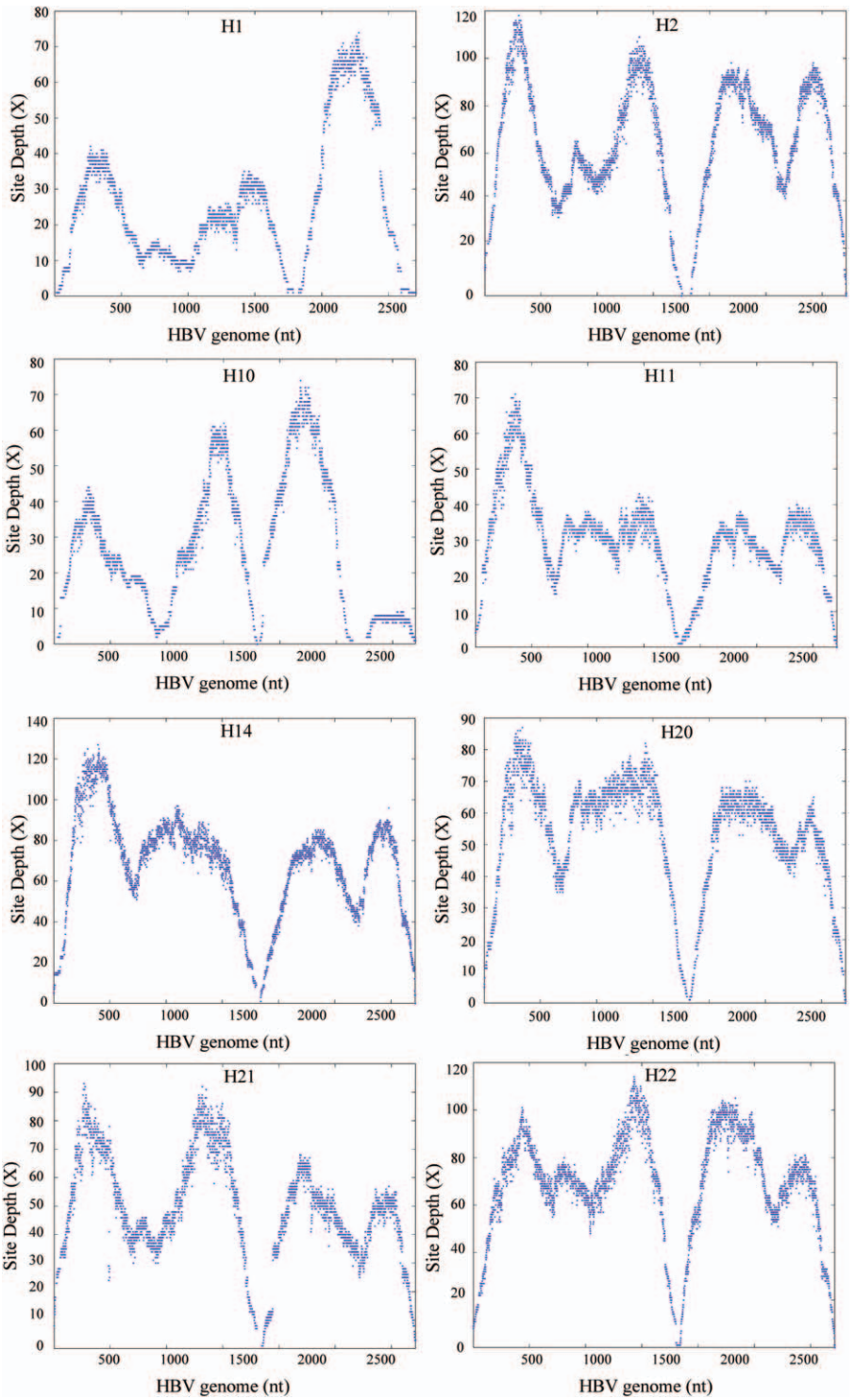

Figure 3. The degree of coverage and depth for all samples. The value of horizontal axis indicates the position of HBV genome. The value of vertical axis indicates the site depth of HBV genome by HTS.

primers resulted in a high failure rate, especially in the low viral load sample. In our preliminary experiments, HBV genome was failed to amplify in 6 out of 7 plasma samples with low viral load HBV ($<10^3$) using a single pair of primers. Surprisingly, regardless of the viral loads, the HBV genome was successfully amplified in all the samples using 2 pairs of nested primers. In agreement with other studies,[23] although amplification of the whole HBV genome using a single pair of primers is convenient and cost-effective, 2 pairs of nested primers turned out to be an effective approach to amplify the whole genome of HBV, especially in low viral load sample. In other studies, however, the HBV genome was divided into 4 or 6 fragments for amplification.[18,20] Comparison of 2 pairs of nested PCR with 4- or 6-pair PCR, we found that the HBV whole genome could be amplified

Table 4

Overview of the iSNV detected in 6 samples, including the frequency and depth of mutations, and the alteration of amino acid position.

Sample	Position	Reference	Mutation	Frequency	AA	Depth	AA position
H1	336	C	T	0.59	Synonymous	32	RT-69
H1	381	G	T	0.45	Synonymous	33	RT-84
H1	531	C	T	0.33	Synonymous	18	RT-134
H1	1290	T	C	0.41	Synonymous	22	
H1	1542	G	A	0.21	G>R	28	
H1	1613	G	A	0.45	R>K	20	
H1	2439	A	G	0.62	E>G	34	
H1	2444	C	A and G	0.35 and 0.03	Q>K and Q>E	31	
H2	378	T	C	0.22	Synonymous	92	RT-83
H2	390	G	A	0.8	Synonymous	79	RT-87
H2	530	A	G and T	0.31 and 0.02	D>G and D>V	49	RT-134
H2	573	C	T	0.51	Synonymous	51	RT-148
H2	2431	A	G	0.31	N>S	74	
H2	2559	C	A and G	0.74 and 0.02	Q>K and Q>E	62	
H2	2640	C	T	0.22	P>F	49	
H2	3016	A	G	0.26	Q>R	73	
H2	3081	C	T and G	0.76 and 0.04	P>S and P>A	54	
H10	273	A	G and C	0.74 and 0.03	Synonymous and Q>H	38	RT-48
H10	293	G	A	0.47	R>H	34	RT-55
H10	357	T	C	0.17	N>K	36	RT-76
H10	1229	A	G	0.22	Q>R	27	
H10	1230	G	C	0.23	Q>R	26	
H10	1993	C	T	0.12	Synonymous	40	
H10	2080	G	A	0.73	Synonymous	48	
H10	2139	C	T	0.89	A>V	61	
H10	2183	C	A	0.1	L>I	70	
H10	2189	A	T	0.9	I>F	68	
H10	2213	A	G	0.59	I>V	59	
H10	2238	A	C	0.26	E>A	62	
H10	2241	C	T	0.15	T>I	65	
H21	342	C	T	0.75	Synonymous	59	RT-71
H21	468	C	G	0.08	Synonymous	64	RT-113
H21	486	C	A	0.71	Synonymous	21	RT-119
H21	488	G	A	0.61	R>K and R>N	26	RT-120
H21	489	G	A and C	0.65 and 0.05	R>K and R>N	20	RT-120
H21	490	A	T and C	0.24 and 0.04	I>S and I>P	25	RT-121
H21	491	T	C	0.11	I>S and I>P	66	RT-121
H21	494	A	T	0.36	N>I	25	RT-122
H21	853	A	C	0.15	Synonymous	41	RT-241
H21	906	A	T	0.68	Synonymous	34	RT-259
H21	1123	A	C	0.12	S>R	50	RT-332
H21	1423	G	A and C	0.73 and 0.02	V>I and V>L	55	
H21	2288	C	T and A	0.43 and 0.15	S>F and S>Y	47	
H21	2444	C	T	0.16	Q>Stop	38	
H22	1613	G	A	0.12	R>K	73	
H22	1752	A	G	0.48	S>G and I>V	23	
H11	480	T	C	0.28	Synonymous	47	
H11	2386	G	A and C	0.5 and 0.03	R>H and R>P	30	
H11	2420	A	T	0.3	Q>L	33	
H11	2525	A	C	0.2	K>N	25	

AA = amino acid, RT = reverse transcript.

by all these methods. However, the latter approaches are in convenient and cost-ineffective, and may introduce more false-positive mutations due to the use of more primers.

In comparison with the results of amplicon-sequencing and direct sequencing without PCR, we found that amplicon-sequencing offered an enormous advantage in terms of the amount of required data. Direct sequencing resulted in 1.76 million of clean reads but with only a few reads mapped to the HBV genome. In contrast, approximately 80% of amplicon sequences are mapped to the HBV genome, requiring fewer sequences to obtain depth coverage. We speculated that the samples may be contaminated with the human genome DNA in the step of nucleic acid extraction or hemolysis during blood sample preparation. Of importance, although the viral loads of HBV were significantly different in the 8 plasma samples, there was little difference in the numbers of reads obtained nested PCR sequencing. We believe that although the reads that were generated from the nested PCR did not map to the HBV genome at a very high rate, nested PCR appears a remarkable improvement and an acceptable mapping rate, compared with the direct sequencing.[4,24,25] Furthermore, the number of effective reads is likely enhanced by increasing the loading

quantity during nested PCR sequencing. Therefore, we conclude that integrating nested PCR with HTS is a more appropriate method for sequencing of the HBV whole genome in low viral load samples than direct sequencing without PCR.

During antiviral therapy with nucleoside analogues (NAs), the emergence of genomic variation is a critical issue. Although it remains unclear whether genomic variation influences the efficiency of antiviral therapy, reports have suggested that viral genomic alterations contribute to the progression of disease status.[26–28] In our study, a total of 50 iSNVs were identified in 6 samples and 21 out of these 50 iSNVs which occurred in the reverse transcriptase (RT)-coding region. We did not identify any known mutations associated with drug resistance to antiviral therapy with NAs in the present study. Nevertheless, mutations in the RT region should be considered when administering NAs. Furthermore, 14 iSNVs were identified with a mutation frequency <20%, indicating that low frequency mutations were widely prevalent in patients. Notably, 30 iSNVs were found in these low viral load samples, whereas only a few mutations were found in high viral load samples. Importantly, mutations in the RT region appeared in all low viral load samples, whereas there were few mutations in this region among high viral load samples. Previous studies have been reported that the isolates with low HBV loads ($<10^5$ copies/mL) were more frequently mutated in core protein (HBc) than the isolates with higher HBV loads ($>10^5$ copies/mL).[29,30] Furthermore, similar results were found in HIV-1 infections that patients with low viral load (<1000 copies/mL) showed an increased resistance mutation.[31,32] Our findings indicate that the genome mutations are more prevalent in the low viral load samples than the high viral load samples, and that mutations in the RT region may play an enigmatic role during the treatment of HBV infection. Indeed, because the source of the samples was collected from the Southern China, where the HBV genotype is mainly B-type, we did not analyze the results for different HBV genotypes. In fact, we aimed to improve the low frequency mutation detection rate of resistant gene in samples of low HBV loads. These findings should be validated by additional studies, and we plan to collect a large number of low HBV load samples in the future to further investigate the effect of low frequency mutations on the treatment of patients with HBV infection.

In conclusion, this paper presents a convenient and cost-effective method to reveal the full landscape of HBV genomes and mutations, which provides an effective way to amplify and sequence the whole HBV genome, especially in samples with low viral load or in low quality samples (e.g., those ones contaminated with human genome), which are usually not appropriate for constructing libraries and direct HTS without PCR amplification. Moreover, our findings suggest that low frequency mutations of the HBV genome should not be neglected, especially among low viral load samples.

Acknowledgments

We appreciate for the analysis of bioinformatics by Ming Ni. We thank Liyun Zheng and Qiuxiang Liu for their help in the proofreading the manuscript.

References

[1] Yonghao G, Jin X, Jun L, et al. An epidemiological serosurvey of hepatitis B virus shows evidence of declining prevalence due to hepatitis B vaccination in central China. Int J Infect Dis 2015;40: 75–80.

[2] Schädler S, Hildt E. HBV life cycle: entry and morphogenesis. Viruses 2009;1:185–209.

[3] Castelain S, Descamps V, Brochot E, et al., High association of T1858-G1896 precore mutations with impaired base pairing and high hepatitis B virus DNA levels in HBeAg-negative chronically infected patients. *Arch Virol*, 2017, 162, 1913–1920.

[4] Widasari DI, Yano Y, Heriyanto DS, et al. A deep-sequencing method detects drug-resistant mutations in the hepatitis B virus in Indonesians. Intervirology 2014;57:384–92.

[5] Locarnini S, Zoulim F. Molecular genetics of HBV infection. Antivir Ther 2010;15(suppl 3):3–14.

[6] Xu Z, Wu G, Li F, et al. Positive selection signals of hepatitis B virus and their association with disease stages and viral genotypes. Infect Genet Evol 2013;19:176–87.

[7] Mese S, Arikan M, Cakiris A, et al. Role of the line probe assay INNO-LiPA HBV DR and ultradeep pyrosequencing in detecting resistance mutations to nucleoside/nucleotide analogues in viral samples isolated from chronic hepatitis B patients. J Gen Virol 2013;94:2729–38.

[8] Barzon L, Lavezzo E, Costanzi G, et al. Next-generation sequencing technologies in diagnostic virology. J Clin Virol 2013;58:346–50.

[9] Szpara ML, Parsons L, Enquist LW. Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. J Virol 2010;84:5303–13.

[10] Sauvage V, Laperche S, Cheval J, et al. Viral metagenomics applied to blood donors and recipients at high risk for blood-borne infections. Blood Transfus 2016;14:400–7.

[11] Gunther S, Li BC, Miska S, et al. 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 1995;69:5437–44.

[12] Sun S, Zhou H, Zhou B, et al. [Sensitivity and specificity of nested PCR pyrosequencing in hepatitis B virus drug resistance gene testing]. Nan Fang Yi Ke Da Xue Xue Bao 2012;32:610–3.

[13] Vincenti D, Solmone M, Garbuglia AR, et al. A sensitive direct sequencing assay based on nested PCR for the detection of HBV polymerase and surface glycoprotein mutations. J Virol Methods 2009;159:53–7.

[14] Taha SE, El-Hady SA, Ahmed TM, et al. Detection of occult HBV infection by nested PCR assay among chronic hepatitis C patients with and without hepatocellular carcinoma. Egypt J Med Hum Genet 2013;14:353–60.

[15] Chen CR, Yuan Q, Ge SX, et al. [Surveillance for occult HBV infection and HBsAg variants in blood donors]. Bing Du Xue Bao 2009;25: 178–84.

[16] Inuzuka T, Ueda Y, Morimura H, et al. Reactivation from occult HBV carrier status is characterized by low genetic heterogeneity with the wild-type or G1896A variant prevalence. J Hepatol 2014;61: 492–501.

[17] Li F, Zhang D, Li Y, et al. Whole genome characterization of hepatitis B virus quasispecies with massively parallel pyrosequencing. Clin Microbiol Infect 2015;21:280–7.

[18] Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–9.

[19] Ali MM, Hasan F, Ahmad S, et al. Comparative evaluation of INNO-LiPA HBV assay, direct DNA sequencing and subtractive PCR-RFLP for genotyping of clinical HBV isolates. Virol J 2010;7:111.

[20] Chook JB, Teo WL, Ngeow YF, et al. Universal primers for detection and sequencing of hepatitis B virus genomes across genotypes A to G. J Clin Microbiol 2015;53:1831–5.

[21] Jin Y, Wei FL, Ma LN, et al. [Comparison of fluorescence quantitative PCR and direct PCR sequencing in the detection of HBV polymerase RT region]. Zhonghua Gan Zang Bing Za Zhi 2010;18:641–5.

[22] White RR, Quake SR, Curr K:. Digital PCR provides absolute quantitation of viral load for an occult RNA virus. J Virol Methods 2012;179:45–50.

[23] Gunther S, Sommer G, Von Breunig F, et al. Amplification of full-length hepatitis B virus genomes from samples from patients with low levels of viremia: frequency and functional consequences of PCR-introduced mutations. J Clin Microbiol 1998;36:531–8.

[24] Ko SY, Oh HB, Park CW, et al. Analysis of hepatitis B virus drug-resistant mutant haplotypes by ultra-deep pyrosequencing. Clin Microbiol Infect 2012;18:E404–11.

[25] Wang Y, Shan X, Liang Z, et al. Deep sequencing analysis of HBV genotype shift and correlation with antiviral efficiency during adefovir dipivoxil therapy. PLoS ONE 2015;10:e131337.

[26] Gunther S, Fischer L, Pult I, et al. Naturally occurring variants of hepatitis B virus. Adv Virus Res 1999;52:25–137.

[27] Parekh S, Zoulim F, Ahn SH, et al. Genome replication, virion secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. J Virol 2003;77:6601–12.
[28] Tu H, Bonura C, Giannini C, et al. Biological impact of natural COOH-terminal deletions of hepatitis B virus X protein in hepatocellular carcinoma tissues. Cancer Res 2001;61:7803–10.
[29] Sendi H, Mehrabmohseni M, Shahraz S, et al. CTL escape mutations of core protein are more frequent in strains of HBeAg negative patients with low levels of HBV DNA. J Clin Virol 2009;46:259–64.
[30] Yuen MF, Sablon E, Libbrecht E, et al. Significance of viral load, core promoter/precore mutations and specific sequences of polymerase gene in HBV-infected patients on 3-year lamivudine treatment. Antivir Ther 2006;11:779–86.
[31] Prosperi MCF, Mackie N, Giambenedetto SD, et al. Detection of drug resistance mutations at low plasma HIV-1 RNA load in a European multicentre cohort study. J Antimicrob Chemother 2011;66:1886–96.
[32] Mackie NE, Phillips AN, Kaye S, et al. Antiretroviral drug resistance in HIV-1 infected patients with low-level viremia. J Infect Dis 2010;201:1303–7.