Virus Genes (2019) 55:610–618
https://doi.org/10.1007/s11262-019-01689-5

ORIGINAL PAPER

Next generation sequencing identifies baseline viral mutants associated with treatment response to pegylated interferon in HBeAg-positive chronic hepatitis B

Natthaya Chuaypen[1] · Sunchai Payungporn[1] · Kittiyod Poovorawan[2] · Watcharasak Chotiyaputta[3] · Teerha Piratvisuth[4] · Pisit Tangkijvanich[5]

Received: 26 March 2019 / Accepted: 18 July 2019 / Published online: 29 July 2019

Abstract

Current data of hepatitis B virus (HBV) variants associated with treatment outcome identified by next generation sequencing (NGS) are limited. This study was aimed at determining the role of baseline sequence variations in the enhancer II (EnhII), basal core promotor (BCP) and pre-core (PC) regions of HBV genotype C in patients treated with pegylated interferon (PEG-IFN). Patients with HBeAg-positive chronic hepatitis B (CHB) treated with 48-week PEG-IFN were enrolled. Combined response (CR) at week 96 was defined by HBeAg seroconversion plus HBV DNA < 2000 IU/mL and HBsAg < 1000 IU/mL. Pre-treatment viral mutations were characterized by Sanger sequencing and NGS (Miseq Illumina platform). Among 47 patients (32 male, mean age 32.4 years), CR was achieved in 12 (25.5%) individuals. Overall, NGS was superior to Sanger sequencing in detecting mutations (61.7% vs. 38.3%, $P < 0.001$). Based on NGS, the prevalence of T1753V (T1753C/A/G) and A1762T/G1764A variants were significantly lower in responders compared to non-responders (8.3% vs. 51.4%, $P = 0.009$ and 33.3% vs. 68.6%, $P = 0.032$, respectively). No significant difference between groups was found regarding C1653T and G1896A mutants. The absence of T1753V and A1762T/G1764A mutations were factors associated with CR (OR 11.65, 95%CI 1.36–100.16, $P = 0.025$, and OR 4.36, 95%CI 1.08–17.63, $P = 0.039$, respectively). The existence of pre-treatment T1753V, A1762T/G1764A mutations and their combination yielded negative predictive values of 94.7%, 85.7% and 93.8%, respectively. The presence of HBV mutants in the BCP region determined by NGS at baseline was associated with poor treatment outcome in patients with HBeAg-positive CHB receiving PEG-IFN.

Keywords NGS · HBV · Mutations · PEG-IFN · Treatment response

Introduction

Currently, approximately 250 million people worldwide are chronically infected with hepatitis B virus (HBV), which leads to the development of chronic hepatitis B (CHB), cirrhosis and hepatocellular carcinoma (HCC) [1]. HBV, a partially double-stranded DNA virus, exhibits high mutation rates due to its lack of proofreading activity during viral

Edited by Wolfram Gerlich.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11262-019-01689-5) contains supplementary material, which is available to authorized users.

✉ Pisit Tangkijvanich
Pisit.T@chula.ac.th; pisittkvn@yahoo.com

[1] Faculty of Medicine, Center of Excellence in Hepatitis and Liver Cancer, Chulalongkorn University, Bangkok, Thailand

[2] Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

[3] Department of Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand

[4] NKC Institute of Gastroenterology and Hepatology, Prince of Songkla University, Songkhla, Thailand

[5] Department of Biochemistry, Faculty of Medicine, Center of Excellence in Hepatitis and Liver Cancer, Chulalongkorn University, 1873 Rama 4 Road, Pathumwan, Bangkok 10330, Thailand

replication [2]. Consequently, the viral population in the host usually consists of remarkable genetic heterogeneity and persists in the form of quasispecies [2, 3]. Previous data have shown that HBV genotypes, mutations, and quasispecies are linked to disease progression and antiviral treatment response [4]. The most common mutations include G to A mutation in the pre-core (PC) region (G1896A) and double mutations in the basal core promoter (BCP, A1762T and G1764A), which result in the impairment of HBeAg production [5]. Moreover, other variants including T1753C/A/G (T1753V) in the BCP and C1653T in the enhancer II (EnhII) region have been increasingly recognized as factors associated with outcome of chronic HBV infection [6]. Previous study reported that A1762T/G1764A mutations were associated with treatment outcome of pegylated interferon (PEG-IFN) [7], suggesting that certain HBV mutations at baseline might affect therapeutic outcome. Despite these findings, several minor mutations of less than 20% are usually not detectable by conventional sequencing methods [8].

Next generation sequencing (NGS) has recently overcome this technical inadequacy and established reliable quantitative assays to verify minor mutations in HBV genome [9, 10]. For instance, NGS of the whole HBV genome revealed the distribution of substitution frequencies and their characteristics in quasispecies that correlated with liver disease progression and severity [8]. In addition, another report of NGS demonstrated a higher sequence diversity was observed in patients with HBeAg-negative CHB than those with HBeAg-positive CHB [11]. Among patients undergoing nucleos(t)ide analogue (NA) therapy, it was recently shown that baseline PC and BCP mutations were associated with reduced likelihood of HBsAg clearance in patients with HBeAg-positive CHB [12]. Moreover, the lack of BCP and PC variants was associated with increased response to PEG-IFN-based therapy in patients with HBeAg-negative CHB [13]. These results indicate that HBV heterogeneity at baseline could be predictive of treatment outcome in patients with CHB. Currently, data on the characterization of HBV quasispecies in patients treated with PEG-IFN are limited. Thus, this study was aimed at investigating the role of baseline viral mutations in the EnhII/BCP/PC regions by NGS in patients receiving PEG-IFN.

Materials and methods

Patients

Between January 2010 and October 2015, 130 patients with HBeAg-positive CHB were treated and completed a full course of 48-week PEG-IFN monotherapy at the King Chulalongkorn Memorial Hospital, Bangkok, Thailand. The indications for treatment were based on the European Association for the Study of the Liver (EASL) guideline. Among them, 47 patients infected with genotype C and had available pre-treatment stored sera were enrolled in this study. All these patients exhibited HBsAg positivity and had elevated serum alanine aminotransferase (ALT) levels and serum HBV DNA levels > 20,000 IU/mL prior to treatment for at least 6 months. In addition, all patients were not co-infected with hepatitis C virus and/or human immunodeficiency virus. The therapeutic endpoint in this report was combined response (CR), defined by HBeAg seroconversion plus HBV DNA < 2000 IU/mL and HBsAg < 1000 IU/mL at week 96. Serum samples collected from each patient at baseline were stored at − 80 °C until further analysis.

All subjects had provided written informed consent as approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University, which followed the Helsinki Declaration and Good Clinical Practice guidelines.

Serological and virological assays

Qualitative HBsAg, HBeAg, and anti-HBe were measured by commercially available enzyme-linked immunosorbent assay kits (Abbott Laboratories, Chicago, IL, USA). HBsAg titers were quantified by the Elecsys HBsAg II Quant reagent kits (Roche Diagnostics, Indianapolis, IN, USA). HBV DNA levels were quantified by the Abbott Real-Time HBV assay (Abbott Laboratories, Chicago, IL, USA), with the lower limit of detection of 10 IU/mL.

Sanger sequencing

To identify mutations in the EnhII/BCP/PC regions, 200 µL of baseline serum samples collected from each patient were extracted for DNA by phenol/chloroform/isoamyl alcohol extraction and eluted in 30 µL of distilled water. The mutations were then amplified using specific primers and identified by direct sequencing [14]. In addition, HBV genotyping was performed based on sequencing and phylogenetic analysis of the pre-S region, as described previously [15].

Next generation sequencing

To identify mutations in the EnhII/BCP/PC regions by NGS, the extracted DNA was amplified by Phusion™High-Fidelity DNA Polymerase; F-530 (Thermo Scientific, Waltham, Massachusetts, USA) with design primer set, located in the conserved region of EnhII/BCP/PC in HBV genome (genotype C). Product size was 448 bp (forward primer: nucleotide position (nt) 1551–1569: GTCTGTGCCTTCTCATCTG and reverse primer: nt 1980–1998: GTGTCGAGGAGATCT CGAA). The thermocycling conditions used were: 98 °C for 30 s and 35 cycles of 98 °C for 10 s, 53 °C for 30 s, and 72 °C for 30 s. Final extension was performed at 72 °C for

7 min. All fragments were amplified under the same conditions using PCR. The PCR products were then processed on 2% of agarose gel and purified by HiyieldTM Gel/PCR DNA fragments extraction kit (RBC Bioscience, New Taipei City, Taiwan). The concentration of DNA was measured by Qubit® Fluorometer and reported in term of ng/μL. DNA library preparation was applied by NEBNext® UltraTM DNA Library Prep Kit for Illumina® (NEB). The PCR products were processed by dual-index paired-end sequencing on Miseq Illumina Next Generation Sequencing according to the manufacture's procedure and were then analyzed by CLC genomic workbench 10.1.1 program. The overlapping pair reads were aligned and merged then low-quality data (Q score < 30) and the adaptor sequence were trimmed. The obtained total potential reads (Q score ≥ 30) and passed filter (PF) reads were aligned and mapped with the EnhII/BCP/PC regions of HBV reference nucleotide sequences (Accession No. AB981580 in the NCBI database; http://www.ncbi.nlm.nih.gov/). The sequencing reads within the EnhII/BCP/PC regions obtained from each sample were clustered into operational taxonomic units (OTUs) and then analyzed for the alpha diversity-Shannon entropy by using Microbial Genomics Module implemented in the CLC Genomics Workbench version 10.1.1 (http://www.clcbio.com/). The nucleotide variations were detected and calculated for percentage of mutations using basic variance detection [8, 9]. In this study, wild-type (WT) strains were classified based on the mutation rate ≤ 1.0%.

Statistical analysis

All analyses were conducted using the SPSS version 22 software (SPSS, Chicago, IL, USA). The results were expressed as mean values ± standard deviation (SD) and percentages. Categorical variables were calculated for statistical significance with the χ^2 test and Fisher's exact test. Comparisons between groups (responders and non-responders) were analyzed by the Mann–Whitney U test for quantitative variables. The logistic regression analysis was used to calculate odd ratio (OR) and 95% confident intervals (CI) relating variables associated with CR. All statistical tests were two-sided with P values < 0.05 considered as statistical significance.

Results

Baseline characteristics of patients

Among 47 patients (32 male, mean age 32.4 years) recruited in this report, CR and HBsAg clearance were achieved in 12 (25.5%) and 4 (8.5%) patients, respectively. Patients' baseline characteristics with respect to CR, non-responders and HBsAg clearance are shown in Table 1. There was no difference between groups regarding mean age, gender distribution, white blood cell and platelet count, serum ALT, HBV DNA, HBsAg levels and liver stiffness measured by transient elastography (FibroScan) at baseline.

NGS improved the detection of mutants compared with Sanger sequencing

The frequencies of EnhII/BCP/PC variants identified by Sanger sequencing and NGS in the same patients were compared. The variants detected by both methods were concordant but NGS had significantly increased the sensitivity for detecting mutants compared with Sanger sequencing (Table 2). An average of 123,591 mapped reads (range 43,441–235,037) and 26,839 coverage (range 8124–118,463) in the EnhII/BCP/PC region for each subject was obtained from NGS and summarized in Supplementary Table 1.

Sanger sequencing was able to detect overall mutants in 38.3% (18/47), whereas NGS could identify the overall variants in 61.7% (29/47) of patients (P < 0.001). Specifically, NGS provided a significant improvement in detecting BCP variants (T1753V, A1762T, and G1764A) and PC variants (G1896A and G1899A). For variants in the EnhII region, NGS also identified higher frequency of C1653T than

Table 1 Baseline characteristics of patients in relation to treatment response

Characteristics	Total (n = 47)	Combined responders (n = 12)	Non-responders (n = 35)	HBsAg clearance (n = 4)	P
Age, years	32.4 ± 9.4	33.4 ± 12.8	32.0 ± 8.1	35.0 ± 22.1	0.842
Gender, % male	32 (68.1)	7 (58.3)	25 (71.4)	3(75.0)	0.297
WBC count, 10^3/ul	6.0 ± 1.6	6.3 ± 1.8	6.0 ± 1.5	6.3 ± 1.2	0.939
Platelet count, 10^3/ul	214.7 ± 60.5	214.4 ± 73.2	214.8 ± 56.6	225.0 ± 234.7	0.874
ALT, U/L	92.5 ± 65.5	78.8 ± 48.5	97.2 ± 70.4	108.8 ± 76.9	0.613
Log_{10} HBV DNA, IU/mL	7.3 ± 1.1	7.6 ± 0.8	7.2 ± 1.1	8.1 ± 0.2	0.280
Log_{10} HBsAg, IU/mL	4.1 ± 0.7	4.2 ± 0.8	4.0 ± 0.7	4.8 ± 0.1	0.123
Liver stiffness, kPa	7.1 ± 3.7	6.1 ± 3.1	7.4 ± 3.9	5.1 ± 2.0	0.499

Table 2 Comparison of NGS and Sanger sequencing for detecting HBV mutations

Nucleotide mutation site	NGS ($N=47$)	Direct sequencing ($N=47$)	P value
EnhII			
C1653T			
Wild type	38 (80.9%)	44 (93.6%)	0.064
Mutant	9 (19.1%)	3 (10.4%)	
BCP			
T1753C/A/G			
Wild type	28 (59.6)	42 (89.4)	0.001*
Mutant	19 (40.4)	5 (10.6)	
A1762T			
Wild type	19 (40.4)	30 (63.8)	0.023*
Mutant	28 (59.6)	17 (36.2)	
G1764A			
Wild type	18 (38.3)	31 (66.0)	0.013*
Mutant	29 (61.7)	16 (34.0)	
A1762T + G1764A			
Wild type	19 (40.4)	30 (63.8)	0.023*
Mutant	28 (59.6)	17 (36.2)	
PC			
G1896A			
Wild type	38 (80.9)	46 (97.9)	0.007*
Mutant	9 (19.1)	1 (2.1)	
G1899A			
Wild type	47 (100%)	47 (100%)	
Mutant	–	–	

Detection number; N (%), *$P<0.05$

Sanger sequencing, although the difference did not reach statistical significance.

Nucleotide diversity in the EnhII/BCP/PC region

To identify the HBV quasispecies in this region, the nucleotide sequences of responders were compared with those of non-responders by Shannon's Entropy. At baseline, responders and non-responders showed comparable nucleotide diversity, as there was no statistical difference in Shannon's Entropy (3.62 vs. 3.65; $P=0.958$) (Fig. 1). The Shannon's Entropy of each sample is also summarized in Supplementary Table 1.

Mutation frequencies varied between responders and non-responders

Based on NGS, there were 19.1% (9/47) patients harboring C1653T, 40.4% (19/47) with T1753V, 59.6% (28/47) with

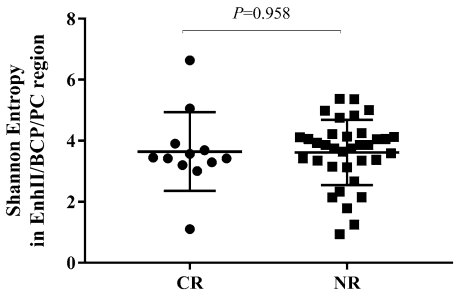

Fig. 1 Scatter plot of Shannon's entropy in the EnhII/BCP/PC regions in relation to treatment response

double BCP mutations and 19.1% (9/47) with G1896A mutant. None of participants had G1899A. Figure 2 demonstrates the distribution of baseline EnhII/BCP/PC mutations detected by NGS in patients who achieved CR (responders, Fig. 2a) and non-responders (Fig. 2b). Overall, 66.7% (8/12) and 22.9% (8/35) of responders and non-responders exhibited WT strains ($P=0.006$), respectively. The percentage of EnhII/BCP/PC mutations detected in each patient is shown in Supplementary Table 2.

Besides the above well-known variants, we examined other hotspot mutations across the EnhII/BCP/PC regions (Supplementary Fig. 1). Regarding treatment response, the frequencies of G1742A, A1752G, G1757A, G1758T, and T1768A variants were significantly higher detected in non-responders compared with responders.

We then compared the prevalence of baseline T1753V and A1762T/G1764A mutations detected in responders and non-responders. To simplify their distribution, the percentage of mutation rates were subcategorized into 4 ranges including 0–0.9% (WT), 1–19%, 20–49% and ≥ 50%, respectively, as shown in Fig. 3. Notably, most of responders exhibited WT at T1753V region as 91.7% (11/12) and 8.3% (1/12) harbored 1–19% mutations (Fig. 3a). For A1762T/G1764A mutations, the majority of responders (66.7%, 8/12) exhibited WT, while 8.3% (1/12), 8.3% (1/12) and 16.7% (2/12) harbored 1–19%, 20–49% and ≥ 50% mutations, respectively (Fig. 3b).

Of note, all 4 patients achieving HBsAg clearance displayed only pre-treatment WT (4/4, 100%) of T1753V, while 75% (3/4) and 25% (1/4) of these patients harbored WT and 1–19% of A1762T/G1764A mutations, respectively.

It should be mentioned that there was no statistically significant in the frequency of T1753V variant between responders and non-responders [0/12 (0%) vs. 5/35 (14.3%), $P=0.191$) based on Sanger sequencing. Similarly, the difference in double BCP mutations was not observed between responders and non-responders [3/12 (25%) vs. 14/35 (40%), $P=0.351$) according to Sanger sequencing.

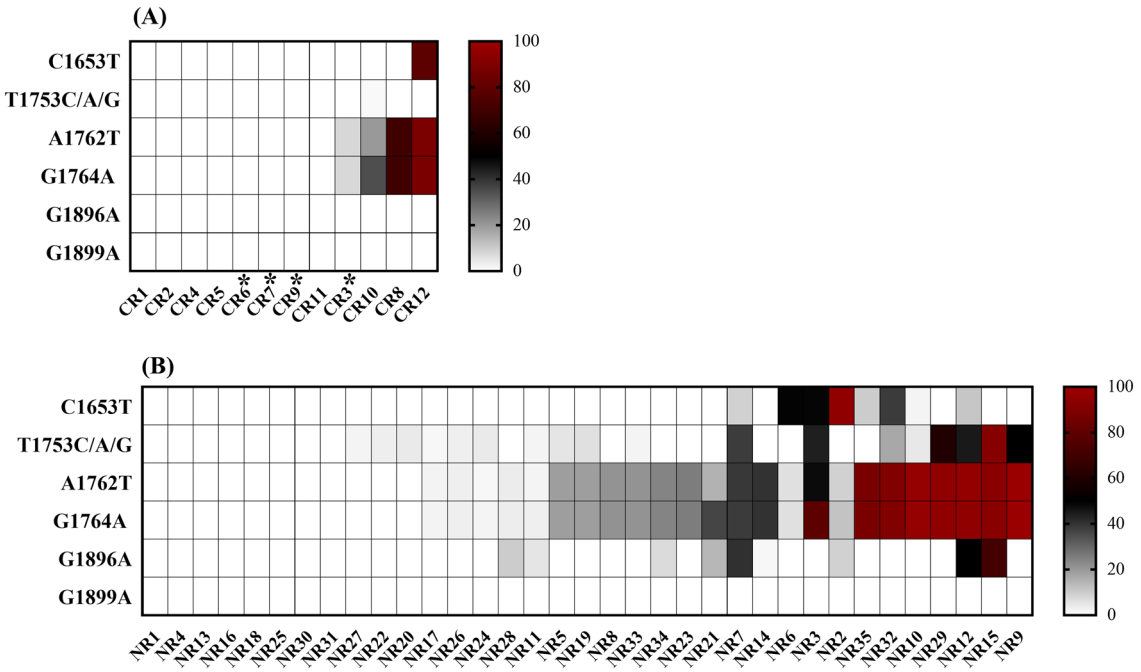

Fig. 2 The distribution of baseline EnhII/BCP/PC mutations detected by NGS **a** responders (CR, *with HBsAg clearance) **b** non-responders (NR)

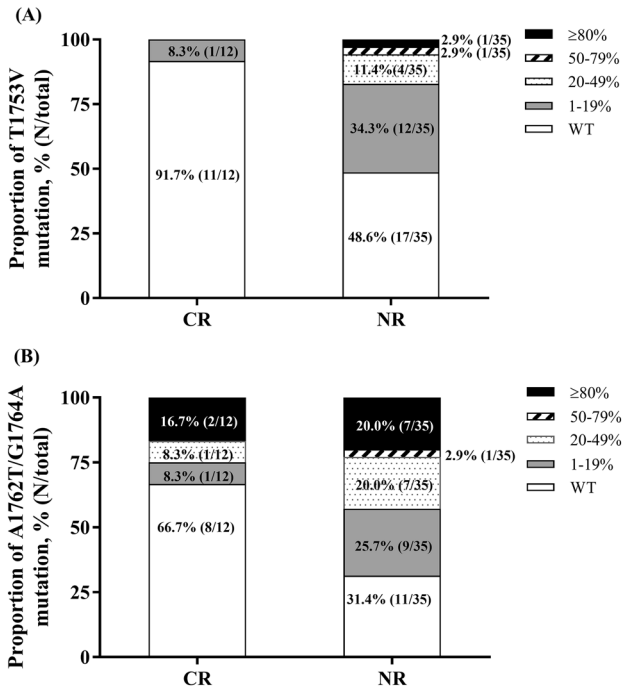

Fig. 3 The proportion of baseline mutations at **a** T1753V and **b** A1762T/G1764A regions detected by NGS in responders (CR) and non-responders (NR)

HBsAg kinetics in relation to viral mutations

The kinetics of HBsAg quantification during and after treatment in relation to baseline T1753V and A1762T/G1764A mutations identified by NGS were further analyzed. Overall, patients with WT showed more decline of HBsAg levels from baseline than patients with mutants (Supplementary Fig. 2).

Baseline predictors of treatment response

Univariate regression analysis was used to identify variables that might be associated with CR. Among baseline characteristics including age, gender, serum ALT, HBV DNA, HBsAg level, and HBV mutations, univariate analysis demonstrated that the absence of T1753V and A1762T/G1764A variants detected by NGS at baseline were associated with CR (Table 3). The predictive values of these mutations at baseline are shown in Table 4. Of noted, the presence of T1753V, A1762T/G1764A variants and their combination exhibited high negative predictive values (NPV) of 94.7%, 85.7%, and 93.8%, respectively.

Discussion

The optimal outcome of achieving HBsAg clearance with undetectable viremia is associated with a decreased risk of developing cirrhosis and HCC [16]. Nonetheless, this

Table 3 Logistic regression analysis of parameters to predict treatment response

Factors	Category	Combined response		HBsAg clearance	
		Univariate analysis		Univariate analysis	
		OR (95%CI)	*P*	OR (95%CI)	*P*
Baseline					
Age, years	< 35 vs. ≥ 35	2.29 (0.43–12.27)	0.333	0.86 (0.08–9.11)	0.901
Gender (% males)	Female vs. male	1.79 (0.46–6.97)	0.404	1.45 (0.14–15.21)	0.758
ALT, U/L	< 80 vs. ≥ 80	1.33 (0.34–5.29)	0.682	1.69 (0.22–13.18)	0.618
Log_{10} HBV DNA, IU/mL	< 7.0 vs. ≥ 7.0	1.77 (0.41–7.75)	0.447	–	–
Log_{10} HBsAg, IU/mL	< 4.0 vs. ≥ 4.0	1.50 (0.38–5.93)	0.563	–	–
Liver stiffness, kPa	< 7.0 vs. ≥ 7.0	1.78 (0.40–7.88)	0.449	0.54 (0.05–5.66)	0.609
Mutations detected by NGS					
C1653T	Wild type vs mutant	3.26 (0.36–29.24)	0.291	–	–
T1753C/A/G		11.65 (1.36–100.16)	0.025*	–	–
A1762T/G1764A		4.36 (1.08–17.63)	0.039*	–	–
G1896A		–	–	–	–
G1899A		–	–	–	–
Mutations detected by Sanger					
C1653T	Wild type vs mutant	0.67 (0.06–8.09)	0.750	–	–
T1753C/A/G		–	–	–	–
A1762T/G1764A		2.00 (0.46–8.71)	0.356	–	–
G1896A		–	–	–	–
G1899A		–	–	–	–

Table 4 Predictive values of mutations for combined response

Factors	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Presence of baseline mutations				
T1753V	91.7	51.4	39.3	94.7
A1762T/G1764A	66.7	68.6	42.1	85.7
Any T1753V or A1764T/G1764A	66.7	77.1	50.0	87.1
Both T1753V and A1764T/G1764A	91.7	42.9	35.5	93.8

PPV positive predictive value, *NPV* negative predictive value

so-called functional cure could be found in a minority of patients with HBeAg-positive CHB receiving PEG-IFN [16]. A more realistic therapeutic endpoint is achieving HBeAg clearance and sustained HBV DNA suppression. However, long-term prognosis of patients with low viral load is still variable depending on HBsAg quantification. For instance, HBsAg < 1000 IU/mL is associated with decreased adverse outcomes in Asian patients with low viremia [17]. In addition, the combined use of HBV DNA (< 2000 IU/mL) and HBsAg (< 1000 IU/mL) exhibited high diagnostic accuracy for predicting inactive carriers with low risk of disease progression [18, 19]. This observation of low serum HBsAg levels may reflect declined intrahepatic covalently closed circular (ccc) DNA and consequently predict a high chance of HBsAg clearance [20]. In this context, we therefore applied combined criteria using sustained viral suppression (HBeAg clearance with HBV DNA < 2000 IU/mL) and HBsAg quantification (< 1000 IU/mL) as the therapeutic endpoint (CR) after 48 weeks of treatment-free follow-up. Based on this endpoint, our data showed that 25.5% and 8.5% of patients achieved CR and HBsAg clearance, respectively.

HBV genotypes B and C are the two most common genotypes distributed in Asia, including Thailand [14]. Current evidence has shown that HBV genotype C, compared with genotype B, is associated with more advanced liver fibrosis and an increased risk of HCC [5]. Moreover, HBV genotype C is more frequently found in HBeAg-positive CHB and displays less favorable response to PEG-IFN [5]. We thus adopted to select only serum samples of patients infected with genotypes C in this cohort to minimize the confounding effect of HBV genotype. Moreover, we characterized HBV quasispecies in the EnhII/BCP/PC regions by NGS.

Compared with direct sequencing, NGS provided a fast and more sensitive quantitative platform for detecting HBV mutational patterns within a detectable range between 1% and 100. It is thus more suitable for detecting minor variants that exist with very low frequencies [10].

Previous studies using direct sequencing had generated conflicting data regarding the predictive role of A1762T/G1764A mutations in determination of PEG-IFN efficiency. For example, the presence of A1762T/G1764A variants before treatment initiation was shown to predict HBeAg clearance after IFN-based treatment [21–23]. In contrast, a recent cohort with a larger sample size of diverse HBV genotypes demonstrated that detectable double BCP mutants at baseline by direct sequencing limited response to PEG-IFN therapy [7]. Our results confirmed and extended the above-mentioned study that BCP variants had a negative influence on PEG-IFN response. Specifically, the presence of A1762T/G1764A identified by NGS at baseline significantly decreased the likelihood of CR in patients infected with HBV genotype C. These results were also in agreement with a recent report of patients receiving NA therapy, which demonstrated that detectable pre-treatment BCP and/or PC mutants by NGS analysis was associated with reduced likelihood of HBsAg clearance [12].

Interestingly, our data showed that the presence of T1753V, another common variant in the BCP region, was predictive of PEG-IFN therapy, whereas C1653T and G1896A mutants were not associated with CR. Whether these mutants might be significantly associated with treatment outcome remain to be evaluated in a higher number of patients. Regarding T1753V, these variants were previously linked to severe acute exacerbation and hepatic decompensation [24]. In a recent meta-analysis, patients harboring T1753V mutations alone and in combination with other variants were more susceptible to acute-on-chronic liver failure [25]. Moreover, the mutants were reported to be associated with progressive liver disease including HCC development in HBV genotype C [26–28]. For example, a recent study from South Korea demonstrated that the presence of T1753V mutations was an independent risk factor for HCC development in long-term follow-up of patients with CHB [28]. Interesting, the occurrence of BCP double mutations alone was not significantly associated with HCC in that report. Together, these data indicated that T1753V mutations might be predictive of developing advanced liver disease in patients infected with HBV genotype C. To our knowledge, there is no previous publication indicating the association of these specific variants with treatment outcome. Of noted, the majority of patients with detectable T1753V in this cohort exhibited low mutation rates below the detecting threshold by direct sequencing (< 20%). Such finding might be a possible explanation that these minor variants existing in the quasispecies pool could have escaped identification in other reports using sequencing methods. Moreover, this observation highlights the importance of minor viral strains associated with reduced treatment response to PEG-IFN, even at frequencies as low as 1% at baseline.

It should be mentioned that the presence of T1753V variants in this cohort usually coexisted along with the occurrence of A1762T/G1764A mutants, but with a relatively lower frequency rates of detection. This finding has confirmed prior data that T1753V variants appear later after the double BCP mutants in the natural course of HBeAg-positive CHB [29]. Notably, our data also revealed that T1753V variants at baseline were more predictive of CR than that of the double BCP mutants. For instance, the NPVs of T1753V variants at baseline for subsequent achieving CR and HBsAg clearance were approximately 95% and 100%, respectively, compared with 85% and 62.8% of the double BCP mutants. Together, it seems that T1753V variants might represent a better biomarker, while A1762T/G1764A mutants also play an important, albeit lesser, role in predicting response to PEG-IFN.

There were some limitations in this study. First, the number of patients was relatively small, particularly those who achieved CR or HBsAg clearance. Also, there was a lack of a validated cohort in our report. In this regard, further large-scale studies are necessary to confirm our observations. Second, we included only HBV genotypes C; thus, it remains to be determined if the results demonstrated here are applicable to other HBV genotypes. Third, whole HBV genome sequencing was not performed in this study. Thus, it remains unknown whether variants in other regions might be associated with treatment outcome. Finally, we identified HBV mutations at baseline, but did not evaluate sequential variants during or after PEG-IFN treatment. Despite this limitation, recent data based on NGS observed no significant dynamic change in nucleotide composition during combined PEG-IFN and NA therapy [13]. As NGS technology has increasingly become faster and less expensive, it is expected that the identification of minor viral strains by NGS-based testing will be applicable in clinical practice. Consequently, NGS platform will offer an advantage of individualized decision-making before initiating PEG-IFN therapy in patients with CHB.

In summary, our results showed that NGS was superior to Sanger sequencing in detecting minor viral strains in the EnhII/BCP/PC regions. Moreover, patients with detectable T1753V and/or A1762T/G1764A variants by NGS exhibited a lower probability of achieving treatment response after PEG-IFN therapy. Considering viral mutations occur gradually through time, our findings suggest that patients with HBeAg-positive CHB should be considered to receive early antiviral treatment before the emerging of the BCP mutants. Nonetheless, further validations in other populations with other HBV genotypes are required.

Acknowledgements This study was funded by the Grant for Chula Research Scholar (CU-GRS-60-06-30-03) and Postdoctoral Fellowship under Rachadapisek Sompot Fund, Chulalongkorn University. The study was also supported by Center of Excellence in Hepatitis and Liver Cancer, Chulalongkorn University, The Thailand Research Fund (RTA5980008) Senior Research Scholar and the Thai Association for the Study of the Liver (THASL).

Author contributions NC performed the experiments, analyzed statistical data, wrote the first draft of the manuscript. SP was involved in experimental design and data analysis. KP, WC, and TP collaborated in the clinical data analysis and reviewed draft manuscript. PT developed experimental design, main analysis plan, and edited manuscript prior to submission. All co-author read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The other authors declare no conflicts of interest.

Ethical approval All subjects had provided written informed consent as approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University (IRB no 016/59), which followed the Helsinki Declaration and Good Clinical Practice guidelines.

References

1. Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ (2015) Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. Lancet 386(10003):1546–1555. https://doi.org/10.1016/S0140-6736(15)61412-X
2. Valaydon ZS, Locarnini SA (2017) The virological aspects of hepatitis B. Best Pract Res Clin Gastroenterol 31(3):257–264. https://doi.org/10.1016/j.bpg.2017.04.013
3. Jazayeri SM, Alavian SM, Carman WF (2010) Hepatitis B virus: origin and evolution. J Viral Hepat 17(4):229–235. https://doi.org/10.1111/j.1365-2893.2009.01193.x
4. Domingo E, Gomez J (2007) Quasispecies and its impact on viral hepatitis. Virus Res 127(2):131–150. https://doi.org/10.1016/j.virusres.2007.02.001
5. Rajoriya N, Combet C, Zoulim F, Janssen HLA (2017) How viral genetic variants and genotypes influence disease and treatment outcome of chronic hepatitis B. Time for an individualised approach? J Hepatol 67(6):1281–1297. https://doi.org/10.1016/j.jhep.2017.07.011
6. Chotiyaputta W, Lok AS (2009) Hepatitis B virus variants. Nat Rev Gastroenterol Hepatol 6(8):453–462. https://doi.org/10.1038/nrgastro.2009.107
7. Sonneveld MJ, Rijckborst V, Zeuzem S, Heathcote EJ, Simon K, Senturk H, Pas SD, Hansen BE, Janssen HL (2012) Presence of precore and core promoter mutants limits the probability of response to peginterferon in hepatitis B e antigen-positive chronic hepatitis B. Hepatology 56(1):67–75. https://doi.org/10.1002/hep.25636
8. Li F, Zhang D, Li Y, Jiang D, Luo S, Du N, Chen W, Deng L, Zeng C (2015) Whole genome characterization of hepatitis B virus quasispecies with massively parallel pyrosequencing. Clin Microbiol Infect 21(3):280–287. https://doi.org/10.1016/j.cmi.2014.10.007
9. Yan L, Zhang H, Ma H, Liu D, Li W, Kang Y, Yang R, Wang J, He G, Xie X, Wang H, Wei L, Lu Z, Shao Q, Chen H (2015) Deep sequencing of hepatitis B virus basal core promoter and precore mutants in HBeAg-positive chronic hepatitis B patients. Sci Rep 5:17950. https://doi.org/10.1038/srep17950
10. Han Y, Gong L, Sheng J, Liu F, Li XH, Chen L, Yu DM, Gong QM, Hao P, Zhang XX (2015) Prediction of virological response by pretreatment hepatitis B virus reverse transcriptase quasispecies heterogeneity: the advantage of using next-generation sequencing. Clin Microbiol Infect. https://doi.org/10.1016/j.cmi.2015.03.021
11. Homs M, Caballero A, Gregori J, Tabernero D, Quer J, Nieto L, Esteban R, Buti M, Rodriguez-Frias F (2014) Clinical application of estimating hepatitis B virus quasispecies complexity by massive sequencing: correlation between natural evolution and on-treatment evolution. PLoS ONE 9(11):e112306. https://doi.org/10.1371/journal.pone.0112306
12. Bayliss J, Yuen L, Rosenberg G, Wong D, Littlejohn M, Jackson K, Gaggar A, Kitrinos KM, Subramanian GM, Marcellin P, Buti M, Janssen HLA, Gane E, Sozzi V, Colledge D, Hammond R, Edwards R, Locarnini S, Thompson A, Revill PA (2017) Deep sequencing shows that HBV basal core promoter and precore variants reduce the likelihood of HBsAg loss following tenofovir disoproxil fumarate therapy in HBeAg-positive chronic hepatitis B. Gut 66(11):2013–2023. https://doi.org/10.1136/gutjnl-2015-309300
13. Jansen L, Welkers MRA, van Dort KA, Takkenberg RB, Lopatin U, Zaaijer HL, de Jong MD, Reesink HW, Kootstra NA (2017) Viral minority variants in the core promoter and precore region identified by deep sequencing are associated with response to peginterferon and adefovir in HBeAg negative chronic hepatitis B patients. Antivir Res 145:87–95. https://doi.org/10.1016/j.antiviral.2017.07.013
14. Tangkijvanich P, Sa-Nguanmoo P, Mahachai V, Theamboonlers A, Poovorawan Y (2010) A case-control study on sequence variations in the enhancer II/core promoter/precore and X genes of hepatitis B virus in patients with hepatocellular carcinoma. Hepatol Int 4(3):577–584. https://doi.org/10.1007/s12072-010-9197-z
15. Sa-Nguanmoo P, Tangkijvanich P, Tharmaphornpilas P, Rasdjarmrearnsook AO, Plianpanich S, Thawornsuk N, Theamboonlers A, Poovorawan Y (2012) Molecular analysis of hepatitis B virus associated with vaccine failure in infants and mothers: a case-control study in Thailand. J Med Virol 84(8):1177–1185. https://doi.org/10.1002/jmv.23260
16. Testoni B, Levrero M, Zoulim F (2017) Challenges to a cure for HBV infection. Semin Liver Dis 37(3):231–242. https://doi.org/10.1055/s-0037-1606212
17. Tseng TC, Liu CJ, Yang HC, Su TH, Wang CC, Chen CL, Hsu CA, Kuo SF, Liu CH, Chen PJ, Chen DS, Kao JH (2013) Serum hepatitis B surface antigen levels help predict disease progression in patients with low hepatitis B virus loads. Hepatology 57(2):441–450. https://doi.org/10.1002/hep.26041
18. Brunetto MR, Oliveri F, Colombatto P, Moriconi F, Ciccorossi P, Coco B, Romagnoli V, Cherubini B, Moscato G, Maina AM, Cavallone D, Bonino F (2010) Hepatitis B surface antigen serum levels help to distinguish active from inactive hepatitis B virus genotype D carriers. Gastroenterology 139(2):483–490. https://doi.org/10.1053/j.gastro.2010.04.052
19. Liu J, Yang HI, Lee MH, Jen CL, Batrla-Utermann R, Lu SN, Wang LY, You SL, Chen CJ (2016) Serum levels of hepatitis B surface antigen and DNA can predict inactive carriers with low risk of disease progression. Hepatology 64(2):381–389. https://doi.org/10.1002/hep.28552
20. Cornberg M, Wong VW, Locarnini S, Brunetto M, Janssen HLA, Chan HL (2017) The role of quantitative hepatitis B surface antigen revisited. J Hepatol 66(2):398–411. https://doi.org/10.1016/j.jhep.2016.08.009
21. Erhardt A, Reineke U, Blondin D, Gerlich WH, Adams O, Heintges T, Niederau C, Haussinger D (2000) Mutations of the

core promoter and response to interferon treatment in chronic replicative hepatitis B. Hepatology 31(3):716–725. https://doi.org/10.1002/hep.510310323
22. Marrone A, Zampino R, Luongo G, Utili R, Karayiannis P, Ruggiero G (2003) Low HBeAg serum levels correlate with the presence of the double A1762T/G1764A core promoter mutation and a positive response to interferon in patients with chronic hepatitis B virus infection. Intervirology 46(4):222–226. https://doi.org/10.1159/000072431
23. Tseng TC, Yu ML, Liu CJ, Lin CL, Huang YW, Hsu CS, Liu CH, Kuo SF, Pan CJ, Yang SS, Su CW, Chen PJ, Chen DS, Kao JH (2011) Effect of host and viral factors on hepatitis B e antigen-positive chronic hepatitis B patients receiving pegylated interferon-alpha-2a therapy. Antivir Ther 16(5):629–637. https://doi.org/10.3851/IMP1841
24. Chen CH, Lin CL, Hu TH, Hung CH, Tseng PL, Wang JH, Chang JY, Lu SN, Chien RN, Lee CM (2014) Entecavir vs. lamivudine in chronic hepatitis B patients with severe acute exacerbation and hepatic decompensation. J Hepatol 60(6):1127–1134. https://doi.org/10.1016/j.jhep.2014.02.013
25. Nian X, Xu Z, Liu Y, Chen J, Li X, Xu D (2016) Association between hepatitis B virus basal core promoter/precore region mutations and the risk of hepatitis B-related acute-on-chronic liver failure in the Chinese population: an updated meta-analysis. Hep Intl 10(4):606–615. https://doi.org/10.1007/s12072-016-9716-7
26. Tanaka Y, Mukaide M, Orito E, Yuen MF, Ito K, Kurbanov F, Sugauchi F, Asahina Y, Izumi N, Kato M, Lai CL, Ueda R, Mizokami M (2006) Specific mutations in enhancer II/core promoter of hepatitis B virus subgenotypes C1/C2 increase the risk of hepatocellular carcinoma. J Hepatol 45(5):646–653. https://doi.org/10.1016/j.jhep.2006.06.018
27. Xu L, Qian G, Tang L, Su J, Wang JS (2010) Genetic variations of hepatitis B virus and serum aflatoxin-lysine adduct on high risk of hepatocellular carcinoma in Southern Guangxi, China. J Hepatol 53(4):671–676. https://doi.org/10.1016/j.jhep.2010.04.032
28. Lee D, Lyu H, Chung YH, Kim JA, Mathews P, Jaffee E, Zheng L, Yu E, Lee YJ, Ryu SH (2016) Genomic change in hepatitis B virus associated with development of hepatocellular carcinoma. World J Gastroenterol 22(23):5393–5399. https://doi.org/10.3748/wjg.v22.i23.5393
29. Takahashi K, Ohta Y, Kanai K, Akahane Y, Iwasa Y, Hino K, Ohno N, Yoshizawa H, Mishiro S (1999) Clinical implications of mutations C-to-T1653 and T-to-C/A/G1753 of hepatitis B virus genotype C genome in chronic liver disease. Arch Virol 144(7):1299–1308