HBV Drug Resistance Substitutions Existed before the Clinical Approval of Nucleos(t)ide Analogues: A Bioinformatic Analysis by GenBank Data Mining

Xizhan Xu, Kuanhui Xiang, Mingze Su, Yao Li, Wei Ji, Yutang Li, Hui Zhuang * and Tong Li * 🆔

Department of Microbiology and Center of Infectious Disease, School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China; xuxizhan2017@163.com (X.X.); xiangkuanhui174@126.com (K.X.); sumingze88@163.com (M.S.); yao.lee@foxmail.com (Y.L.); jiwei_yunlong@126.com (W.J.); liyutang@bjmu.edu.cn (Y.L.)

* Correspondence: zhuangbmu@126.com (H.Z.); tongli08@vip.sina.com (T.L.); Tel.: +86-010-8280-2413 (T.L.)

Abstract: Naturally occurring nucleos(t)ide analogue resistance (NUCr) substitution frequencies in the reverse transcriptase (RT) of the hepatitis B virus (HBV) were studied extensively after the clinical approval of nucleos(t)ide analogues (NUCs; year of approval 1998). We aimed to study NUCr substitutions in HBV RT sequences obtained before 1998 and better understand the evolution of RT sequences without NUC pressures. Our strategy was to retrieve HBV sequences from GenBank deposited before 1998. The initial search used the keywords "hepatitis B virus" or "HBV" and 1139 sequences were found. Data analyses included information extraction: sequence quality control and amino acid substitution analysis on 8 primary NUCr and 3 secondary substitution codons. Three hundred and ninety-four RT-containing sequences of 8 genotypes from 25 countries in 4 continents were selected. Twenty-seven (6.9%) sequences were found to harbor substitutions at NUCr-related codons. Secondary substitutions (rtL80V and rtV173G/A/L) occurred more frequently than primary NUCr substitutions (rtI169L; rtA181G; T184A/S; rtS202T/R; rtM204L and rtM250K). Typical amino acid substitutions associated with NUCr were of rtL80V, rtV173L and rtT184A/S. We confirm the presence of naturally occurring typical HBV NUCr substitutions with very low frequencies, and secondary substitutions are more likely to occur than primary NUCr substitutions without the selective pressure of NUCs.

Keywords: hepatitis B virus; reverse transcriptase; nucleos(t)ides analogue resistance; substitution; pre-existing; naturally occurring

1. Introduction

Hepatitis B virus (HBV) infection remains a serious public health issue [1]. Six antiviral nucleos(t)ide analogues (NUCs), namely lamivudine (LAM), adefovir dipivoxil (ADV), entecavir (ETV), telbivudine (LdT), tenofovir disoproxil fumarate (TDF) and tenofovir alafenamide fumarate (TAF) have been successively approved for treating chronic hepatitis B (CHB) since 1998 [2–5]. They are all HBV polymerase/reverse transcriptase (RT) inhibitors and can efficiently suppress HBV replication. However, because RT lacks proofreading activity, HBV replication is error-prone and leads to the incorporation of many mutations in its genome [6]. The clinical use of NUCs was limited by the development of NUC resistance (NUCr) substitutions in RT sequences during long-term therapy until the emergence of ETV, TDF and TAF, which have high genetic barriers to resistance [2,7]. However, pre-existing NUCr substitutions can be detected in treatment-naïve patients, which have been shown to affect the antiviral responses [8–10].

According to several clinical practice guidelines and authoritative reviews, NUCr substitutions can be classified into two categories: primary NUCr substitutions at 8 codons (rtI169T, rtA181T/V, rtT184A/C/F/G/I/L/M/S, rtA194T, rtS202C/G/I, rtM204I/V, rtN236T and rtM250I/L/V) and secondary substitutions at 3 codons in RT (rtL80I/V, rtV173L and rtL180M), which are often found in treated patients and termed as classic amino acid (AA) substitutions [2,4,11–15]. Primary NUCr substitutions are known to reduce HBV susceptibility to NUCs but with the cost of impaired viral replication, which have been well characterized in vitro [11]. Secondary substitutions contribute to restoring the functional defects of RT caused by primary NUCr substitutions [14,16]. The relationship between AA substitutions at these codons and NUCr was extensively studied after 1998 [2,4]. Resistance to LAM or LdT is conferred by rtM204I/V, which is often accompanied by rtL80I/V, rtL180M and/or rtV173L. Resistance to ADV is linked to rtA181T/V or rtN236T. Substitution of rtA181T/V can also cause decreased susceptibility to LAM or LdT. Resistance to ETV needs a combination of substitutions of rtM204I/V and rtL180M plus one of the following substitutions: rtI169T, rtT184A/C/F/G/I/L/M/S, rtS202C/G/I or rtM250I/L/V. Resistance to TDF/TAF has not been well elucidated yet [2,4,11]. LAM, the first approved oral NUC in 1998, has been excluded by several clinical practice guidelines as a first-line antiviral agent due to the high frequency of NUCr substitution. ETV and TDF/TAF have the lowest rates of NUCr and are recommended as the first-line antiviral NUCs for CHB today [5,12,13].

In addition, through the review of the published literature and a systematic analysis of potential substitutions associated with NUCr, we also summarized some non-classic substitutions in the RT domain of HBV polymerase that appeared sporadically in clinical practice and might contribute to decreasing susceptibility to NUCs or improve the virus fitness (Table 1). For example, Hayashi et al. [17] characterized two novel substitutions, rtI163V and rtA186T, in an ETV-refractory patient. The rtA186T substitution was responsible for ETV resistance and significantly reduced viral replication efficacy, while rtI163V substitution could sustain viral fitness and rescue the variant. Shirvani-Dastgerdi et al. [18] identified rtS78T as being partially responsible for multidrug resistance to ETV and TDF in two patients receiving ETV-TDF combination therapy. However, the emergence of NUCr caused by these non-classic substitutions remains limited in clinical practice.

Previous studies have reported that HBV NUCr substitutions might pre-exist in the RT region in treatment-naïve patients [8–10]. They might come from naturally occurring substitutions or horizontal transmission from NUC-exposed patients. However, the pre-existing NUCr substitution frequencies reported varied from 0% to 35.7% in previous studies [8–11,19–24]. For instance, our group has reported that no pre-existing NUCr substitution was observed among 192 treatment-naïve Chinese patients using a population-sequencing approach, while Tan et al. found a relatively high incidence of rtM204I/V substitution (23.3%) in 1042 untreated CHB patients [10,11]. According to a recent meta-analysis of 106 studies involving 12212 treatment-naïve CHB patients, the summarized incidence of NUCr substitution frequency worldwide was 5.73% (95% confidence interval: 4.85–6.61%), which was obviously lower than the abnormally high prevalence of NUCr given by Tan et al. [10,24]. We speculate that such a wide range of NUCr substitution frequencies in treatment-naïve patients might be due to discrepancies of the following factors: study design, sample sizes/sources, substitution testing methods, and so on [20,25]. More importantly, several key factors can largely affect the results, leading to inappropriate interpretation of the true prevalence of pre-existing NUCr substitutions. First, patients enrolled in some studies might receive antiviral therapy without awareness of the exposure, i.e., lack of critical validation for treatment-naïvety at enrollment [20]. Second, treatment-naïve patients might be infected with NUCr HBV strains through horizontal transmission. Third, different studies might focus on investigating inconsistent NUCr substitution codons. For example, initial NUCr studies often investigated rtM204I/V substitution, while more AA substitution codons associated with NUCr in the RT domain have been included in more recent studies [10,26]. Characterizing NUCr substitutions in patients before the clinical approval of NUCs is a good way to illustrate if the pre-existing NUCr

substitutions are naturally occurring. However, the serum samples collected before 1998 are difficult to obtain.

Table 1. Summary of potential AA substitutions associated with NUCr in treated CHB patients.

Drugs	Primary NUCr Substitutions	Secondary Substitutions
LAM	**rtA181T/V** [2,4,11], rtA181S [27], **rtM204I/V** [2,4,11], rtM204Q/S [28,29]	rtH55R [30], **rtL80I/V** [2,4,11], rtS117F [31], rtT128N [32], rtW153Q [32], **rtV173L** [2,4,11], **rtL180M** [2,4,11], rtL180C [33], rtA200V [34], rtV207I [35], rtL217P [36], rtL229F [37], rtQ267H [38], rtL269I [39]
ADV	rtS78T [40], **rtA181T/V** [2,4,11], rtA181S [41], rtE218G [42], rtI233V [43], **rtN236T** [2,4,11], rtN236V [44]	rtL80I/V [45], rtV84M [46], rtS85A [46], rtV214A [47], rtL217R [48], rtP237H [47]
ETV	rtS78T [18], **rtI169T** [2,4,11], **rtA181T/V** [2,4,11], **rtT184A/C/F/G/I/L/M/S** [2,4,11], rtA186T [17], **rtS202C/G/I** [2,4,11], **rtM204I/V** [2,4,11], **rtM250I/V/L** [2,4,11]	rtI163V [17], rtS219A [49], rtY245H [49], rtS246T [50], rtL269I [39]
LdT	**rtM204I/V** [2,4,11], **rtA181T/V** [2,4,11]	
TDF/TAF	rtS78T [18], rtP177G [51], rtA181T/V [2,4,11], **rtA194T** [2,4,11], rtF249A [51]	

Classic NUCr AA substitutions recommended by several clinical practice guidelines and authoritative reviews are indicated in bold [2,4,5,11–13]. Plain text represents non-classic AA substitutions that occurred sporadically. AA: amino acid; ADV: adefovir dipivoxil; CHB: chronic hepatitis B; ETV: entecavir; LAM: lamivudine; LdT: telbivudine; NUCr: nucleos(t)ide analogue resistance; rt: reverse transcriptase; TDF: tenofovir disoproxil fumarate; TAF: tenofovir alafenamide fumarate.

In this study, we analyzed HBV RT sequences before the clinical approval of NUCs to characterize the pre-existing NUCr substitutions by using publically available GenBank data. Our study focuses on analyzing the classic NUCr substitutions recommended by several clinical practice guidelines. These results can provide critical information regarding pre-existing (largely due to naturally occurring) NUCr substitutions in HBV RT sequences and a better understanding of HBV evolution without selective pressure of NUCs.

2. Materials and Methods

2.1. Sequence Search and Qualification Strategy

We performed a GenBank database (https://www.ncbi.nlm.nih.gov/nuccore) search using keywords "hepatitis B virus" or "HBV" to retrieve sequences deposited before December 1998, the year just before the clinical approval of NUCs. Each sequence with a unique accession number was retained after manually removing the redundant sequences by checking duplicate sequence accession numbers. The following sequences were then excluded: (1) artificial synthetic constructs; (2) chromosomally integrated HBV sequences with human genes; (3) non-human HBV sequences; (4) other types of hepatitis virus sequences (hepatitis A virus, hepatitis C virus, hepatitis D virus or hepatitis E virus); (5) expression vector sequences; (6) sequences from patients co-infected with HBV/human immunodeficiency virus (HIV) receiving antiviral therapy against HIV; (7) sequences from CHB patients who participated in the clinical trials (phase I, II and III) for LAM; and (8) HBV sequences without the RT region. After exclusion, the rest of the sequences containing the HBV RT region were qualified for further analysis.

2.2. Data Extraction

The selection of each sequence and data extraction were reviewed and performed independently by two authors (X.X. and K.X.). For each qualified sequence, the following data were extracted: the GenBank accession number, related nucleotide sequence, sequence release year, sample origin, and sequence length. References related to specific sequences were thoroughly analyzed to collect the relevant information if the GenBank records did not provide it.

2.3. HBV Genotyping and Detection of Recombinant Sequences

HBV genotypes were identified using the National Center for Biotechnology Information (NCBI) viral genotyping tool (http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi). The results were reconfirmed using the online software tool HBV STAR (http://www.vgb.ucl.ac.uk/starn.shtml) [52]. An online tool jpHMM (http://jphmm.gobics.de/submission_HBV) designed for viruses with circular genomes was used to predict genotype recombination [53].

2.4. Systematic Analysis of Potential NUCr Substitutions

Potential AA substitutions in RT domain of HBV polymerase associated with NUCr were listed in Table 1. This table was modified from Menéndez-Arias et al. [4] and Liu et al. [11], and updated according to the latest literature. All these AA substitutions have been reported as being selected during NUCs therapy and verified by in vitro phenotypic data.

2.5. HBV Drug Resistance Substitution Analysis

HBV RT sequences were extracted from the qualified HBV sequences, which contained various lengths of the RT region. Multiple RT nucleotide sequences were aligned using BioEdit7.0 software (Ibis Biosciences, Carlsbad, CA, USA). All RT nucleotide sequences were translated in frame to AA sequences. Each genetic codon harboring degenerate base was manually translated to the corresponding AA. Classic substitutions at 11 codons that have been often suggested for monitoring NUCr in treated patients were analyzed in this study. The obtained AA sequences were compared with the consensus RT AA sequences to identify the NUCr substitutions at rtL80, rtI169, rtV173, rtL180, rtA181, rtT184, rtA194, rtS202, rtM204, rtN236 and rtM250 [2,54–56]. The codons that were mixed with wild-type and substituted AAs were recorded as indicating the presence of a substitution, which might reflect the presence of HBV quasispecies [11,54]. All sequences with NUCr substitutions were further determined by an HBV drug resistance interpretation online tool (http://www.hiv-grade.de/HBV_grade/deployed/grade.pl?program=HBValg).

2.6. Statistical Analysis

Statistical analysis was performed using SPSS 20.0 software (IBM, Armonk, NY, USA). Chi-square tests were used for categorical data. p Values were calculated with the two-tailed method. p Value < 0.05 was considered statistically significant.

3. Results

3.1. Selection for RT-Containing HBV Sequences

A total of 1139 sequences with unique accession number were found from the GenBank nucleotide database using the keywords "hepatitis B virus" or "HBV" before December 1998. The sequence exclusion and inclusion strategy and processing were illustrated by the flowchart in Figure 1. Seven hundred and eight HBV sequences were identified, in which 314 sequences were excluded because they did not harbor an RT region. Finally, 394 RT-containing (full-length: 134; partial-length: 260) sequences were included for further analysis. Background information (i.e., accession numbers, release time,

origin, RT coverage and sequence length) for all qualified sequences is presented in Supplementary Materials (Table S1).
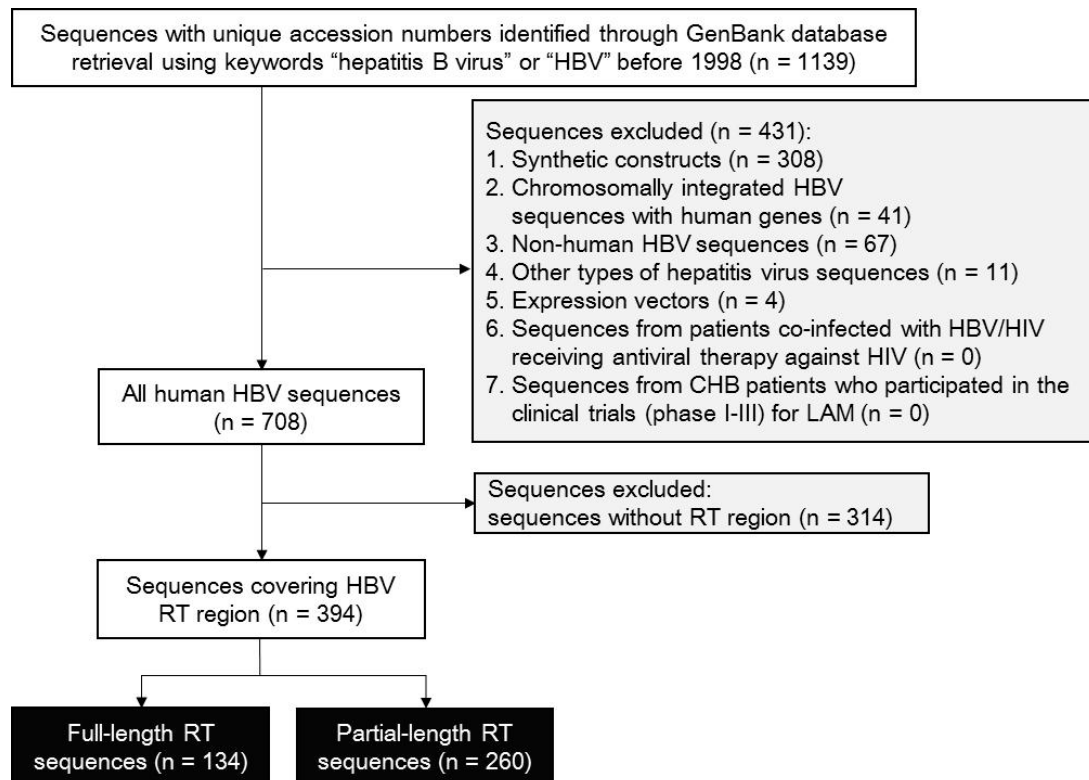

Figure 1. Flowchart showing the process of searching and screening for HBV RT sequences used in this study. The grey boxes show the excluded sequences, while the black boxes show those that qualified for analyses. CHB: chronic hepatitis B; HBV: hepatitis B virus; HIV: human immunodeficiency virus; LAM: lamivudine; RT, reverse transcriptase.

3.2. Quality Control for HBV RT-Containing Sequences

The selected RT sequences were further evaluated in terms of their coverage of the studied NUCr-related codons. One hundred and thirty-four sequences were full-length RT sequences, which contained all 11 codons of interest. The other 260 partial-length RT sequences ranged from 23 to 305 AAs in length with a median length of 226 AAs. Figure 2A illustrates the overlapping features of the recruited RT sequences. Our data suggested that most of previous studies appeared to primarily focus on HBV small surface proteins, especially on the 'a' determinant region, which is overlapped by the RT region. The RT sequences were further grouped according to their coverage, as shown in Figure 2B. In total, 90.1% (355/394) of the recruited RT sequences carried 5–11 analyzed NUCr substitution codons and only 28 sequences (7.1%) did not cover any studied NUCr codons.

Due to the diversity of the sequences in length and their coverage of the studied codons, the substitution analysis strategy was codon-based rather than sequence-based. The numbers of the available AAs at 11 codons in RT were summarized in Figure 2C. Thus, the substitution frequency was calculated as the percentage of the numbers of substituted AAs with respect to the available AAs at this codon.
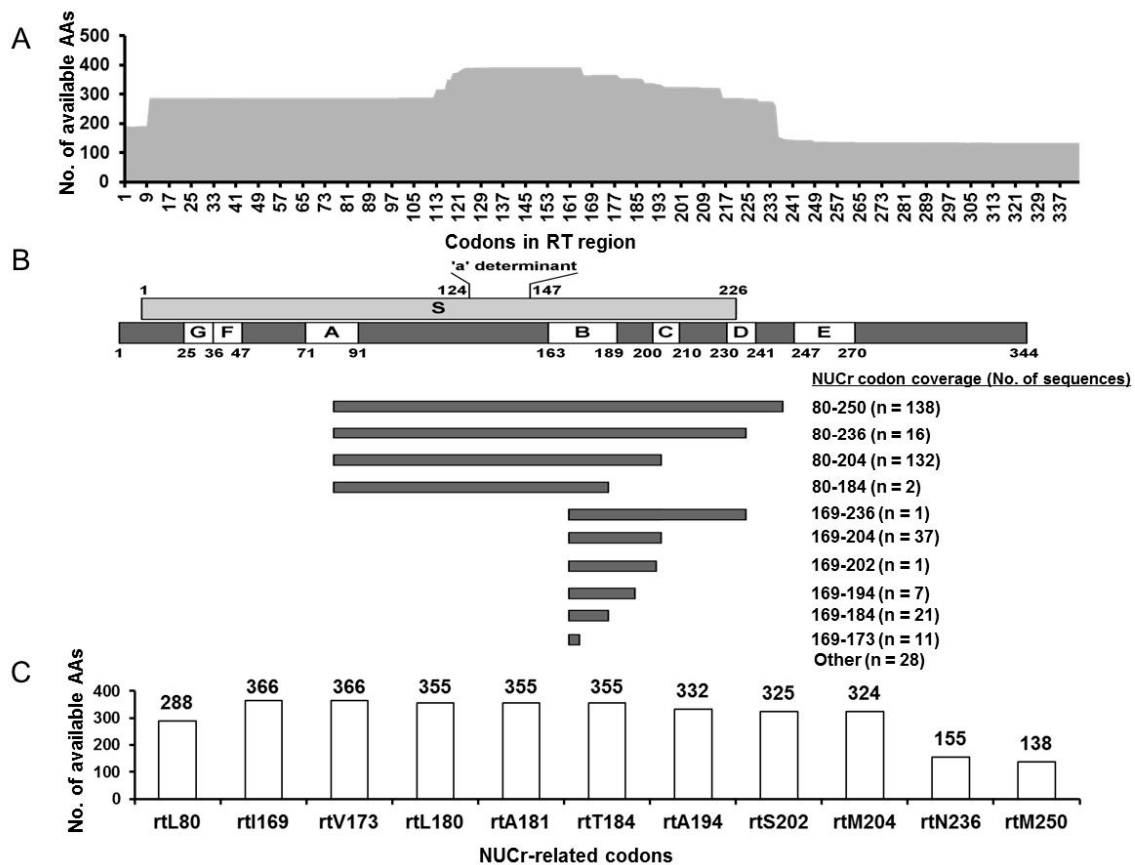
Figure 2. Schematic illustration of the analyzed RT sequences. (**A**) The total number of available AAs at each codon in 394 HBV RT sequences; (**B**) All sequences are clustered based on the different coverage of analyzed NUCr-related codons. For example, "80–250" shows where these sequences include 11 studied NUCr-related codons. "Other" shows where no NUCr-related codons are covered by these sequences; (**C**) The number of available AAs for each of the NUCr-related codons studied is summarized. AA: amino acid; HBV: hepatitis B virus; NUCr: nucleos(t)ide analogue resistance; RT: reverse transcriptase.

3.3. RT Sequence Origins and HBV Genotypes

The finally analyzed 394 HBV RT sequences were found from 25 countries within 4 continents. The sequences from Africa, America, Asia and Europe accounted for 4.8%, 23.1%, 41.9% and 30.2%, respectively. Eight HBV genotypes (A: 27.4%; B: 9.4%; C: 31.0%; D: 16.8% E: 2.3%; F: 10.7%; G: 0.5%; H: 0.5%) were found (Figure S1). Moreover, four recombinant sequences were found, which did not affect the NUCr-related codons and could be used for substitution analysis (Figure S2). The geographical distribution of HBV genotypes was consistent with previous reports [57]. Notably, although the dominant genotype in America was F (35.2%), these sequences were mainly from Central and South America (Argentina), not from North America.

3.4. NUCr Substitution Analysis

Among the 394 sequences, 27 (6.9%) were found to harbor substitutions at NUCr-related codons (Table 2). Eleven sequences had substitutions at 6 codons associated with primary NUCr, and 16 carried secondary substitutions at 2 codons. However, no combination substitutions at multiple codons were found. No substitution at rtA194, rtN236, or rtL180 was detected.

Table 2. The AA substitutions identified at the primary NUCr and secondary substitution codons in RT-containing sequences.

Substitution Category	NUCr-Related Codons	Substitution Types	No. of Substituted AAs	No. of Available AAs	Substitution Frequency (%) [1]
Primary	rtI169	L	2	366	0.5
NUCr	rtA181	G	1	355	0.3
substitution	rtT184	A [3], S [3]	1, 1	355	0.6
	rtA194	ND	0	332	0
	rtS202	T, R	2, 1	325	0.9
	rtM204	L	2	324	0.6
	rtN236	ND	0	155	0
	rtM250	K	1	138	0.7
Subtotal	8 codons	6 codons, 8 types	11	2350	0.5 [2]
Secondary	rtL80	V [3]	7	288	2.4
substitution	rtV173	G, A, L [3]	5, 2, 2	366	2.5
	rtL180	ND	0	355	0
Subtotal	3 codons	2 codons, 4 types	16	1008	1.6 [2]
Total	11 codons	8 codons, 12 types	27	3358	0.8

[1] The substitution frequency at a studied codon was calculated as the percentage of the numbers of substituted AAs with respect to the available AAs at this codon; [2] The substitution frequency was significantly higher for the secondary substitution codons (16/1008, 1.6%) as compared to the primary NUCr substitution codons (11/2350, 0.5%) (χ^2 = 11.08, *p* = 0.0009). Chi-square test was used; [3] AA substitution types belonged to the well-known typical NUCr substitutions. AA: amino acid; ND: not detected; NUCr: nucleos(t)ide analogue resistance; RT: reverse transcriptase.

Moreover, the substitution frequency was significantly higher at the secondary substitution codons (16/1008, 1.6%) than that at the primary NUCr codons (11/2350, 0.5%) (χ^2 = 11.08, *p* = 0.0009). Among eight codons with detectable substitutions, rtV173 had the highest substitution frequency (9/366, 2.5%), which was significantly higher than the overall substitution frequency (27/3358, 0.8%) (χ^2 = 9.44, *p* = 0.0021). Interestingly, only rtL80V (*n* = 7), rtV173L (*n* = 2) and rtT184A/S (*n* = 2) were typical NUCr substitutions, whose ability to confer resistance to NUCs have been well elucidated in vitro [2,4,11]. However, the relationship between other atypical AA substitutions with susceptibility to NUCs has not been characterized in vitro, such as rtI169L (not typical rtI169T) and rtM204L (not typical rtM204I/V). Thus, the substitution frequency for typical AA substitutions at NUCr-related codons was as low as 0.3% (11/3358).

3.5. Characterization of the RT Sequences Harboring AA Substitutions Associated with NUCr

Table 3 provides the detailed characteristics of 27 RT sequences with AA substitutions at NUCr-related codons. Four sequences (14.8%) belonged to genotype A, 1 (3.7%) belonged to B, 16 (59.3%) belonged to C, and 6 (22.2%) belonged to D. Overall, the NUCr substitutions occurred in 3.7% (4/108) of A, 2.7% (1/37) of B, 13.1% (16/122) of C and 9.1% (6/66) of D genotype sequences. No substitutions at these codons were detected in HBV E to H and recombinant genotypes. The detection frequencies of NUCr substitutions were significantly different among genotypes A to D (*p* = 0.0375). This was mainly due to the significant difference between genotype A (3.7%) and C (13.1%) (*p* = 0.0172). Comparisons between the other two genotypes, i.e., A vs. B, A vs. D, B vs. C, B vs. D and C vs. D, did not show any significant differences (*p* > 0.05).

As shown in Table 3, these 27 sequences were distributed in 12 countries in 4 continents. The substitution frequencies were 5.3% (1/19) for Africa, 2.2% (2/91) for America, 10.3% (17/165) for Asia and 5.9% (7/119) for Europe, respectively (*p* = 0.0930). Asia contributed 63.0% (17/27) of sequences with substitutions at NUCr-related codons, in which 6 and 6 were from China and Japan, respectively (Figure S3).

Table 3. Characteristics of the RT sequences with NUCr-related AA substitutions.

Accession Number	Releasing Year	Sample Origin	Sequencing Method	Sequence Length (bp)	RT Coverage (AA)	Genotype	Substitution	References
X01587.1	1983	Japan	Clone	3214	1–344	C	V173G	[58]
M23808.1	1985	Japan	Clone	681	10–235	C	V173G	[59]
E01164.1	1986	Japan	Clone	748	6–253	C	V173G	NA
M38636.1	1988	South Korea	Clone	3213	1–344	C	V173G	[60]
U19777.1	1988	China	Clone	731	1–243	C	T184A	NA
M27765.1	1989	Japan	Clone	678	10–234	C	I169L	[61]
X14193.1	1989	South Korea	Clone	3213	1–344	C	V173G	[62]
S50225.1	1992	USA	Clone	3222	1–344	A	V173A	[63]
S56208.1	1992	India	Clone	123	169–208	D	S202T	[64]
X77310.1	1994	Italy	Direct	1401	1–300	D	S202R	NA
X80925.1	1995	UK	Direct	3182	1–344	D	S202T	[65]
D50519.1	1996	Japan	Clone	3215	1–344	C	A181G	[66]
AF013631.1	1997	China	Clone	681	10–235	C	L80V	[67]
AF036236.1	1997	China	Clone	838	1–236	C	L80V	NA
AF036237.1	1997	China	Clone	838	1–236	C	L80V	NA
AF036238.1	1997	China	Clone	838	1–236	C	L80V	NA
AF036239.1	1997	China	Clone	838	1–236	C	L80V	NA
U87739.1	1997	South Africa	Direct	846	1–235	A	L80V	[68]
AB014382.1	1998	Japan	Direct	3215	1–344	C	L80V	[69]
AJ003027.1	1998	Germany	Clone	1158	1–235	D	I169L	[70]
AF043564.1	1998	Argentina	Direct	217	122–193	A	V173A	[71]
AJ003026.1	1998	Germany	Clone	1164	1–235	D	V173L	[70]
AJ003028.1	1998	Germany	Clone	1164	1–235	D	V173L	[70]
AF044985.1	1998	Belgium	Clone	297	117–215	B	T184S	NA
AF074449.1	1998	Thailand	Direct	720	10–248	C	M204L	[72]
AF075604.1	1998	Thailand	Direct	720	10–248	C	M204L	NA
AJ012207.1	1998	Germany	Clone	3221	1–344	A	M250K	NA

AA: amino acid; bp: base pair; NA: not available; NUCr: nucleos(t)ide analogue resistance; RT: reverse transcriptase; UK: the United Kingdom of Great Britain and Northern Ireland; USA: the United States of America.

Moreover, the original studies submitting these 27 sequences were traced and carefully studied. We found that 17 sequences were from published papers, while the others were a direct submission to GenBank. Twelve of these 17 sequences were obtained by clone sequencing. These studies aimed to analyze the sequence variations, mainly focusing on HBV surface proteins (M23808 [59], X14193 [62], M27765 [61], S56208 [64], AF013631 [67], AF043564 [71] and AJ003026-AJ003028 [70]). However, none of them carried out in vitro study to verify whether the identified variants were viable or not.

4. Discussion

To the best of our knowledge, this is the first study to analyze HBV NUCr-related substitutions from HBV RT sequences obtained before the clinical approval of NUCs. Our results demonstrate the pre-existence of HBV NUCr substitution, which may largely attribute to the true naturally occurring substitutions and reflect the natural evolution characteristics of HBV variants with NUCr-related substitutions.

The acquisition of high-quality data is critical to a superior GenBank-based data mining study. After a rigorous screening process, 394 HBV RT sequences belonging to genotype A to H from 25 countries in 4 continents were recruited for this study, which showed a wide range of data sources. Moreover, a codon-based rather than a sequence-based method was used to calculate the substitution frequency because qualified HBV RT sequences had different lengths. This strategy could ensure not only an accurate calculation of the substitution frequency at each studied codon, but also an effective use of the available information.

In this study, AA substitutions at NUCr-related codons of RT were found in 72.7% (8/11) of the studied codons with frequencies ranging from 0.3% to 2.5% (Table 2). Secondary substitutions (1.6%) occurred more frequently than primary NUCr substitutions (0.5%). Interestingly, 59.3% (16/27) had an atypical AA substitution type at these codons, meaning that their roles in NUCr and viral fitness are still unknown, as is the case for rtM204L. This phenomenon was more obvious at the primary NUCr substitution codons (81.8%, 9/11) than at the secondary substitution codons (43.8%, 7/16) (Table 2).

Therefore, our findings highly suggested that HBV could naturally mutate at these codons free of NUCs, but those variants carrying typical primary NUCr substitution(s) hardly evolve, probably due to their reduced survival fitness [2].

Our results show that only 11 sequences (2.8%) had typical NUCr-related substitutions. Substitutions rtL80V (*n* = 7) or rtV173L (*n* = 2) are typical secondary substitutions that can restore functional defects of HBV polymerase caused by primary LAM or LdT resistance substitutions [2,4,11]. These naturally occurring secondary substitutions might reduce the resistance barrier to LAM or LdT [2]. Substitution rtT184A/S (*n* = 2) was the only naturally occurring substitution detected as belonging to a typical primary NUCr substitution. It has been related to ETV resistance in the presence of the rtM204V + rtL180M backbone [73]. However, rtT184A/S alone would not lead to clinical resistance to ETV [74,75].

In addition, we found that the detected atypical substitutions of rtI169L, rtA181G, rtS202T/R, rtM204L, and rtM250K were not reported in NUCs-treated patients by the Stanford HBV Drug Resistance Database (https://hivdb.stanford.edu/HBV/HBVseq/development/HBVseq.html; accessing date: 17 January 2017). This suggests that these atypical substitutions might not play important roles in the emergence of NUCr as compared to rtA181T/V, rtS202C/G/I, rtM204I/V and rtM250I/L/V [2,76,77]. Consistent with our findings, atypical substitutions of rtI169L, rtS202R, and rtM204L have been detected in six treatment-naïve CHB patients with low proportions using a high-throughput sequencing method [78]. Kim et al. found that three patients had an rtI169L substitution at a level of 1–8% using a clone sequencing method [79]. We speculate that the reasons for the presence of such atypical substitutions might be due to a quasispecies property of HBV or sequencing errors [80–82]. Moreover, some variants might not be necessarily viable or infectious. Current evidence suggests that these atypical pre-existing variants have a limited role in the development of NUCr. Tenney et al. reported that substitution at rtI169 could only modestly decrease the susceptibility to ETV in the recombinant viruses with LAM resistance and rtM250V substitutions, increasing the viral fitness for survival [74]. Furthermore, in vitro studies showed that substitutions rtS202R and rtM250K would severely impair the replication capability of HBV variants with LAM resistance substitutions (rtL180M + rtM204V), while substitution rtS202T played only a minor role in ETV resistance [77]. Substitution rtM204L was shown to be sensitive to LAM in a study [83]. No study has been performed to elucidate the relationship between substitution rtA181G and ADV resistance.

Except for substitutions detected at eight codons, our data showed no substitutions at rtA194, rtN236 and rtL180. Since there were only 155 AAs at rtN236 available for analysis, this data might not be as solid as for rtA194 and rtL180, each of which had 332 and 355 available AAs, respectively (Figure 2). Moreover, by searching the Stanford HBV Drug Resistance Database, we found that rtA194, rtN236 and rtL180 in NUC treatment-naïve patients had very low substitution frequencies at 0.4% (14/3218), 0.03% (1/3218) and 0.2% (6/3218), respectively [84]. Considering that HBV RT sequences in this database were obtained both before and after 1998, our data together with their data suggested that the true naturally occurring substitutions at these codons should be very low.

We also found that genotype C sequences had the highest substitution frequency (16/122, 13.1%) at NUCr-related codons. However, we should notice that only 7 of 16 sequences harbored typical NUCr AA substitution types, i.e., 6 of rtL80V and 1 of rtT184A. Notably, 4 of 6 rtL80V-containing sequences were from different clones of the same patient. Thus, the naturally occurring NUCr substitutions in different HBV genotypes might not be as different as they seemed to be. This is in line with many previous drug resistance studies carried out after 1998 [2].

Although this study is novel in investigating the occurrence of NUCr substitutions before NUCs were approved, some limitations exist. Firstly, there was a limited number of sequences available in the GenBank database before 1998. Secondly, most sequences were from Asia, Europe and America (375/394, 95.2%) and data from Africa, a highly endemic area of HBV infection, is not adequate. Thirdly, some minor mutants in the quasispecies pool might not be detected due to the lack of a sensitive detection method. In fact, most substitutions (20/27, 74.1%) found in this study were determined by

clone sequencing. Fourth, the possibility of sequencing errors or potential typos while submitting to GenBank could not be completely ruled out, especially for relatively rare AA substitutions, although the original studies submitting these sequences have been carefully reviewed.

In conclusion, we confirmed the presence of naturally occurring typical NUCr substitutions of rtL80V, rtV173L and rtT184A/S with low frequencies before the clinical approval of NUCs. Atypical substitutions of rtI169L, rtA181G, rtS202T/R, rtM204L, rtM250K were also detected. However, more clinical studies are still needed to explore the impacts of these baseline NUCr-related substitutions on antiviral responses.

Supplementary Materials: The following are available online at www.mdpi.com/1999-4915/9/8/199/s1, Figure S1: HBV genotypes identified in this study and their geographical distributions; Figure S2: The HBV sequences with recombinant genotypes found in this study; Figure S3: The distribution of HBV RT sequences with NUCr-related substitutions in different countries; Table S1: Summary of the relevant information of HBV sequences used in this study.

Author Contributions: X.X. carried out data mining, data analysis and wrote the paper; X.X., K.X. and Yao L. worked on substitution analysis; X.X., M.S. and Yutang L. helped with local database setup, data interpretation and figure design; W.J. helped with information extraction from GenBank; H.Z. and T.L. contributed to study design, manuscript preparation and critical review.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Schweitzer, A.; Horn, J.; Mikolajczyk, R.T.; Krause, G.; Ott, J.J. Estimations of worldwide prevalence of chronic hepatitis B virus infection: A systematic review of data published between 1965 and 2013. *Lancet* **2015**, *386*, 1546–1555. [CrossRef]

2. Zoulim, F.; Locarnini, S. Hepatitis B virus resistance to nucleos(t)ide analogues. *Gastroenterology* **2009**, *137*, 1593–1608. [CrossRef] [PubMed]

3. De Clercq, E.; Ferir, G.; Kaptein, S.; Neyts, J. Antiviral treatment of chronic hepatitis B virus (HBV) infections. *Viruses* **2010**, *2*, 1279–1305. [CrossRef] [PubMed]

4. Menendez-Arias, L.; Alvarez, M.; Pacheco, B. Nucleoside/nucleotide analog inhibitors of hepatitis B virus polymerase: Mechanism of action and resistance. *Curr. Opin. Virol.* **2014**, *8*, 1–9. [CrossRef] [PubMed]

5. Hepatol, J. European association for the study of the liver. EASL 2017 clinical practice guidelines on the management of hepatitis B virus infection. *J. Hepatol.* **2017**, *67*, 370–397. [CrossRef]

6. Girones, R.; Miller, R.H. Mutation rate of the hepadnavirus genome. *Virology* **1989**, *170*, 595–597. [CrossRef]

7. Kang, L.; Pan, J.; Wu, J.; Hu, J.; Sun, Q.; Tang, J. Anti-HBV drugs: Progress, unmet needs, and new hope. *Viruses* **2015**, *7*, 4960–4977. [CrossRef] [PubMed]

8. Li, X.; Liu, Y.; Zhao, P.; Wang, Y.; Chen, L.; Xin, S.; Zhang, X.X.; Xu, D. Investigation into drug-resistant mutations of HBV from 845 nucleoside/nucleotide analogue-naïve Chinese patients with chronic HBV infection. *Antivir. Ther.* **2015**, *20*, 141–147. [CrossRef] [PubMed]

9. Han, Y.; Huang, L.H.; Liu, C.M.; Yang, S.; Li, J.; Lin, Z.M.; Kong, X.F.; Yu, D.M.; Zhang, D.H.; Jin, G.D.; et al. Characterization of hepatitis B virus reverse transcriptase sequences in Chinese treatment naïve patients. *J. Gastroenterol. Hepatol.* **2009**, *24*, 1417–1423. [CrossRef] [PubMed]

10. Tan, Y.W.; Ge, G.H.; Zhao, W.; Gan, J.H.; Zhao, Y.; Niu, Z.L.; Zhang, D.J.; Chen, L.; Yu, X.J.; Yang, L.J. YMDD motif mutations in chronic hepatitis B antiviral treatment naïve patients: A multi-center study. *Braz. J. Infect. Dis.* **2012**, *16*, 250–255. [CrossRef]

11. Liu, B.M.; Li, T.; Xu, J.; Li, X.G.; Dong, J.P.; Yan, P.; Yang, J.X.; Yan, L.; Gao, Z.Y.; Li, W.P.; et al. Characterization of potential antiviral resistance mutations in hepatitis B virus reverse transcriptase sequences in treatment-naïve Chinese patients. *Antivir. Res.* **2010**, *85*, 512–519. [CrossRef] [PubMed]

12. Terrault, N.A.; Bzowej, N.H.; Chang, K.M.; Hwang, J.P.; Jonas, M.M.; Murad, M.H. AASLD guidelines for treatment of chronic hepatitis B. *Hepatology* **2016**, *63*, 261–283. [CrossRef] [PubMed]

13. Sarin, S.K.; Kumar, M.; Lau, G.K.; Abbas, Z.; Chan, H.L.; Chen, C.J.; Chen, D.S.; Chen, H.L.; Chen, P.J.; Chien, R.N.; et al. Asian-pacific clinical practice guidelines on the management of hepatitis B: A 2015 update. *Hepatol. Int.* **2016**, *10*, 1–98. [CrossRef] [PubMed]

14. Shaw, T.; Bartholomeusz, A.; Locarnini, S. HBV drug resistance: Mechanisms, detection and interpretation. *J. Hepatol.* **2006**, *44*, 593–606. [CrossRef] [PubMed]

15. World Health Organization. *Guidelines for the Prevention, Care and Treatment of Persons with Chronic Hepatitis B Infection*; World Health Organization: Geneva, Switzerland, 2015.

16. Lok, A.S.; Zoulim, F.; Locarnini, S.; Bartholomeusz, A.; Ghany, M.G.; Pawlotsky, J.M.; Liaw, Y.F.; Mizokami, M.; Kuiken, C. Antiviral drug-resistant HBV: Standardization of nomenclature and assays and recommendations for management. *Hepatology* **2007**, *46*, 254–265. [CrossRef] [PubMed]

17. Hayashi, S.; Murakami, S.; Omagari, K.; Matsui, T.; Iio, E.; Isogawa, M.; Watanabe, T.; Karino, Y.; Tanaka, Y. Characterization of novel entecavir resistance mutations. *J. Hepatol.* **2015**, *63*, 546–553. [CrossRef] [PubMed]

18. Shirvani-Dastgerdi, E.; Winer, B.Y.; Celia-Terrassa, T.; Kang, Y.; Tabernero, D.; Yagmur, E.; Rodriguez-Frias, F.; Gregori, J.; Luedde, T.; Trautwein, C.; et al. Selection of the highly replicative and partially multidrug resistant rtS78T HBV polymerase mutation during TDF-ETV combination therapy. *J. Hepatol.* **2017**, *67*, 246–254. [CrossRef] [PubMed]

19. Nguyen, M.H.; Garcia, R.T.; Trinh, H.N.; Nguyen, H.A.; Nguyen, K.K.; Nguyen, L.H.; Levitt, B. Prevalence of hepatitis B virus DNA polymerase mutations in treatment-naïve patients with chronic hepatitis B. *Aliment. Pharmacol. Ther.* **2009**, *30*, 1150–1158. [CrossRef] [PubMed]

20. Vutien, P.; Trinh, H.N.; Garcia, R.T.; Nguyen, H.A.; Levitt, B.S.; Nguyen, K.; da Silveira, E.; Daugherty, T.; Ahmed, A.; Garcia, G.; et al. Mutations in HBV DNA polymerase associated with nucleos(t)ide resistance are rare in treatment-naïve patients. *Clin. Gastroenterol. Hepatol.* **2014**, *12*, 1363–1370. [CrossRef] [PubMed]

21. Pollicino, T.; Isgro, G.; Di Stefano, R.; Ferraro, D.; Maimone, S.; Brancatelli, S.; Squadrito, G.; Di Marco, V.; Craxi, A.; Raimondo, G. Variability of reverse transcriptase and overlapping S gene in hepatitis B virus isolates from untreated and lamivudine-resistant chronic hepatitis B patients. *Antivir. Ther.* **2009**, *14*, 649–654. [PubMed]

22. Nishijima, N.; Marusawa, H.; Ueda, Y.; Takahashi, K.; Nasu, A.; Osaki, Y.; Kou, T.; Yazumi, S.; Fujiwara, T.; Tsuchiya, S.; 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. [CrossRef] [PubMed]

23. Gomes-Gouvea, M.S.; Ferreira, A.C.; Teixeira, R.; Andrade, J.R.; Ferreira, A.S.; Barros, L.M.; Rezende, R.E.; Nastri, A.C.; Leite, A.G.; Piccoli, L.Z.; et al. HBV carrying drug-resistance mutations in chronically infected treatment-naïve patients. *Antivir. Ther.* **2015**, *20*, 387–395. [CrossRef] [PubMed]

24. Zhang, Q.; Liao, Y.; Cai, B.; Li, Y.; Li, L.; Zhang, J.; An, Y.; Wang, L. Incidence of natural resistance mutations in naïve chronic hepatitis B patients: A systematic review and meta-analysis. *J. Gastroenterol. Hepatol.* **2015**, *30*, 252–261. [CrossRef] [PubMed]

25. Zoulim, F. Hepatitis B virus resistance to entecavir in nucleoside naïve patients: Does it exist? *Hepatology* **2006**, *44*, 1404–1407. [CrossRef] [PubMed]

26. Huang, Z.M.; Huang, Q.W.; Qin, Y.Q.; He, Y.Z.; Qin, H.J.; Zhou, Y.N.; Xu, X.; Huang, M.J. YMDD mutations in patients with chronic hepatitis B untreated with antiviral medicines. *World J. Gastroenterol.* **2005**, *11*, 867–870. [CrossRef] [PubMed]

27. Karatayli, E.; Karayalcin, S.; Karaaslan, H.; Kayhan, H.; Turkyilmaz, A.R.; Sahin, F.; Yurdaydin, C.; Bozdayi, A.M. A novel mutation pattern emerging during lamivudine treatment shows cross-resistance to adefovir dipivoxil treatment. *Antivir. Ther.* **2007**, *12*, 761–768. [PubMed]

28. Liu, Y.; Xu, Z.H.; Wang, Y.; Li, X.D.; Liu, L.M.; Chen, L.; Xin, S.J.; Xu, D.P. rtM204Q may serve as a novel lamivudine-resistance-associated mutation of hepatitis B virus. *PLoS ONE* **2014**, *9*. [CrossRef] [PubMed]

29. Bozdayi, A.M.; Uzunalimoglu, O.; Turkyilmaz, A.R.; Aslan, N.; Sezgin, O.; Sahin, T.; Bozdayi, G.; Cinar, K.; Pai, S.B.; Pai, R.; et al. YSDD: A novel mutation in HBV DNA polymerase confers clinical resistance to lamivudine. *J. Viral Hepat.* **2003**, *10*, 256–265. [CrossRef] [PubMed]

30. Xiang, K.H.; Zhao, C.Y.; Wang, S.; Li, Y.; Su, M.Z.; Wang, Q.Y.; Xu, X.Z.; Deng, J.; Zhuang, H.; Li, T. Impacts of HBV rtH55R polymerase substitution on viral replication and rtM204I/V resistance to nucleoside/nucleotide antiviral drugs. *Antivir. Ther.* **2017**. [CrossRef] [PubMed]

31. Lin, C.L.; Chien, R.N.; Hu, C.C.; Lai, M.W.; Yeh, C.T. Identification of hepatitis B virus rtS117F substitution as a compensatory mutation for rtM204I during lamivudine therapy. *J. Antimicrob. Chemother.* **2012**, *67*, 39–48. [CrossRef] [PubMed]

32. Torresi, J.; Earnest-Silveira, L.; Civitico, G.; Walters, T.E.; Lewin, S.R.; Fyfe, J.; Locarnini, S.A.; Manns, M.; Trautwein, C.; Bock, T.C. Restoration of replication phenotype of lamivudine-resistant hepatitis B virus mutants by compensatory changes in the "fingers" subdomain of the viral polymerase selected as a consequence of mutations in the overlapping S gene. *Virology* **2002**, *299*, 88–99. [CrossRef] [PubMed]

33. Pai, S.B.; Bozdayi, A.M.; Pai, R.B.; Beker, T.; Sarioglu, M.; Turkyilmaz, A.R.; Grier, J.; Yurdaydin, C.; SChinazi, R.F. Emergence of a novel mutation in the FLLA region of hepatitis B virus during lamivudine therapy. *Antimicrob. Agents Chemother.* **2005**, *49*, 2618–2624. [CrossRef] [PubMed]

34. Fu, L.; Cheng, Y.C. Role of additional mutations outside the YMDD motif of hepatitis B virus polymerase in L(-)SddC (3TC) resistance. *Biochem. Pharmacol.* **1998**, *55*, 1567–1572. [CrossRef]

35. Zollner, B.; Sterneck, M.; Wursthorn, K.; Petersen, J.; Schroter, M.; Laufs, R.; Feucht, H.H. Prevalence, incidence, and clinical relevance of the reverse transcriptase V207I mutation outside the YMDD motif of the hepatitis B virus polymerase during lamivudine therapy. *J. Clin. Microbiol.* **2005**, *43*, 2503–2505. [CrossRef] [PubMed]

36. Ji, D.; Liu, Y.; Si, L.L.; Li, L.; Chen, G.F.; Xin, S.J.; Zhao, J.M.; Xu, D. Variable influence of mutational patterns in reverse-transcriptase domain on replication capacity of hepatitis B virus isolates from antiviral-experienced patients. *Clin. Chim. Acta* **2011**, *412*, 305–313. [CrossRef] [PubMed]

37. Ji, D.; Liu, Y.; Li, L.; Xu, Z.H.; Si, L.L.; Dai, J.Z.; Li, X.D.; Wang, L.; Yao, Z.T.; Xin, S.J.; et al. The rtL229 substitutions in the reverse transcriptase region of hepatitis B virus (HBV) polymerase are potentially associated with lamivudine resistance as a compensatory mutation. *J. Clin. Virol.* **2012**, *54*, 66–72. [CrossRef] [PubMed]

38. Qin, B.; Zhang, B.; Zhang, X.D.; He, T.T.; Xu, W.Y.; Fu, L.J.; Tu, C.Y. Substitution rtQ267H of hepatitis B virus increases the weight of replication and lamivudine resistance. *Hepat. Mon.* **2013**, *13*. [CrossRef] [PubMed]

39. Ahn, S.H.; Kim, D.H.; Lee, A.R.; Kim, B.K.; Park, Y.K.; Park, E.S.; Ahn, S.H.; Shin, G.C.; Park, S.; Kang, H.S.; et al. Substitution at rt269 in hepatitis B virus polymerase is a compensatory mutation associated with multi-drug resistance. *PLoS ONE* **2015**, *10*. [CrossRef] [PubMed]

40. Cento, V.; Van Hemert, F.; Neumann-Fraune, M.; Mirabelli, C.; Di Maio, V.C.; Salpini, R.; Bertoli, A.; Micheli, V.; Gubertini, G.; Romano, S.; et al. Anti-HBV treatment induces novel reverse transcriptase mutations with reflective effect on HBV S antigen. *J. Infect.* **2013**, *67*, 303–312. [CrossRef] [PubMed]

41. Liu, Y.; Li, X.; Xin, S.; Xu, Z.; Chen, R.; Yang, J.; Liu, L.; Wong, V.W.; Yang, D.; Chan, H.L.; et al. The rtA181S mutation of hepatitis B virus primarily confers resistance to adefovir dipivoxil. *J. Viral Hepat.* **2015**, *22*, 328–334. [CrossRef] [PubMed]

42. Liu, L.J.; Wang, J.H.; Du, S.C.; Tian, J.H.; Yang, R.F.; Wei, L. rtE218G, a novel hepatitis B virus mutation with resistance to adefovir dipivoxil in patients with chronic hepatitis B. *J. Viral Hepat.* **2010**, *17*, 66–72. [CrossRef] [PubMed]

43. Schildgen, O.; Olotu, C.; Funk, A.; Zollner, B.; Helm, M.; Rockstroh, J.K.; Sirma, H. Selection and counterselection of the rtI233V adefovir resistance mutation during antiviral therapy. *J. Clin. Microbiol.* **2010**, *48*, 631–634. [CrossRef] [PubMed]

44. Liu, Y.; Liu, W.H.; Li, X.D.; Xu, Z.H.; Wang, X.; Li, C.X.; Chen, L.; Xin, S.J.; Xu, D.P. Screening and identification of a novel adefovir dipivoxil resistance associated mutation, rtN236V, of HBV from a large cohort of HBV-infected patients. *Antivir. Ther.* **2014**, *19*, 551–558. [CrossRef] [PubMed]

45. Warner, N.; Locarnini, S.; Kuiper, M.; Bartholomeusz, A.; Ayres, A.; Yuen, L.; Shaw, T. The L80I substitution in the reverse transcriptase domain of the hepatitis B virus polymerase is associated with lamivudine resistance and enhanced viral replication in vitro. *Antimicrob. Agents Chemother.* **2007**, *51*, 2285–2292. [CrossRef] [PubMed]

46. Sheldon, J.; Rodes, B.; Zoulim, F.; Bartholomeusz, A.; Soriano, V. Mutations affecting the replication capacity of the hepatitis B virus. *J. Viral Hepat.* **2006**, *13*, 427–434. [CrossRef] [PubMed]

47. Ghany, M.; Liang, T.J. Drug targets and molecular mechanisms of drug resistance in chronic hepatitis B. *Gastroenterology* **2007**, *132*, 1574–1585. [CrossRef] [PubMed]

48. Bottecchia, M.; Madejon, A.; Sheldon, J.; Garcia-Samaniego, J.; Barreiro, P.; Soriano, V. Hepatitis B virus genotype A2 harbours an L217R polymorphism which may account for a lower response to adefovir. *J. Antimicrob. Chemother.* **2008**, *62*, 626–627. [CrossRef] [PubMed]

49. Karatayli, E.; Karatayli, S.C.; Cinar, K.; Gokahmetoglu, S.; Guven, K.; Idilman, R.; Yurdaydin, C.; Bozdayi, A.M. Molecular characterization of a novel entecavir mutation pattern isolated from a multi-drug refractory patient with chronic hepatitis B infection. *J. Clin. Virol.* **2012**, *53*, 130–134. [CrossRef] [PubMed]

50. Hu, J.L.; Cui, J.; Guo, J.J.; Zhang, W.L.; Cai, X.F.; Yuan, Z.W.; Li, Q.L.; Deng, X.Y.; Zeng, A.Z.; Hu, Y.; et al. Phenotypic assay of a hepatitis B virus strain carrying an rtS246T variant using a new strategy. *J. Med. Virol.* **2012**, *84*, 34–43. [CrossRef] [PubMed]

51. Qin, B.; Budeus, B.; Cao, L.; Wu, C.C.; Wang, Y.; Zhang, X.Y.; Rayner, S.; Hoffmann, D.; Lu, M.J.; Chen, X.W. The amino acid substitutions rtP177G and rtF249A in the reverse transcriptase domain of hepatitis B virus polymerase reduce the susceptibility to tenofovir. *Antivir. Res.* **2013**, *97*, 93–100. [CrossRef] [PubMed]

52. Myers, R.; Clark, C.; Khan, A.; Kellam, P.; Tedder, R. Genotyping hepatitis B virus from whole- and sub-genomic fragments using position-specific scoring matrices in HBV STAR. *J. Gen. Virol.* **2006**, *87*, 1459–1464. [CrossRef] [PubMed]

53. Schultz, A.K.; Bulla, I.; Abdou-Chekaraou, M.; Gordien, E.; Morgenstern, B.; Zoaulim, F.; Deny, P.; Stanke, M. jpHMM: Recombination analysis in viruses with circular genomes such as the hepatitis B virus. *Nucleic Acids Res.* **2012**, *40*, 193–198. [CrossRef] [PubMed]

54. Su, M.; Xiang, K.; Li, Y.; Li, Y.; Deng, J.; Xu, X.; Yan, L.; Zhuang, H.; Li, T. Higher detection rates of amino acid substitutions in HBV reverse transcriptase/surface protein overlapping sequence is correlated with lower serum HBV DNA and HBsAg levels in HBeAg-positive chronic hepatitis B patients with subgenotype B2. *Infect. Genet. Evol.* **2016**, *40*, 275–281. [CrossRef] [PubMed]

55. Yang, J.X.; Liu, B.M.; Li, X.G.; Yan, C.H.; Xu, J.; Sun, X.W.; Wang, Y.H.; Jiao, X.J.; Yan, L.; Dong, J.P.; et al. Profile of HBV antiviral resistance mutations with distinct evolutionary pathways against nucleoside/nucleotide analogue treatment among Chinese chronic hepatitis B patients. *Antivir. Ther.* **2010**, *15*, 1171–1178. [CrossRef] [PubMed]

56. Zoulim, F.; Locarnini, S. Optimal management of chronic hepatitis B patients with treatment failure and antiviral drug resistance. *Liver Int.* **2013**, *33*, 116–124. [CrossRef] [PubMed]

57. Sunbul, M. Hepatitis B virus genotypes: Global distribution and clinical importance. *World J. Gastroenterol.* **2014**, *20*, 5427–5434. [CrossRef] [PubMed]

58. Fujiyama, A.; Miyanohara, A.; Nozaki, C.; Yoneyama, T.; Ohtomo, N.; Matsubara, K. Cloning and structural analyses of hepatitis B virus DNAs, subtype adr. *Nucleic Acids Res.* **1983**, *11*, 4601–4610. [CrossRef] [PubMed]

59. Takeshima, H.; Inokoshi, J.; Namiki, M.; Shimada, J.; Omura, S. Structural analysis of the gene coding for hepatitis B virus surface antigen and its product. *J. Gen. Virol.* **1985**, *66 Pt 1*, 195–200. [CrossRef] [PubMed]

60. Mukaide, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiji, K. The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and phylogenetic analysis. *Nucleic Acids Res.* **1992**, *20*, 6105. [CrossRef] [PubMed]

61. Okamoto, H.; Omi, S.; Wang, Y.; Itoh, Y.; Tsuda, F.; Tanaka, T.; Akahane, Y.; Miyakawa, Y.; Mayumi, M. The loss of subtypic determinants in alleles, d/y or w/r, on hepatitis B surface antigen. *Mol. Immunol.* **1989**, *26*, 197–205. [PubMed]

62. Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S. The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype adr. *Nucleic Acids Res.* **1989**, *17*, 2124. [CrossRef] [PubMed]

63. Wands, J.R.; Liang, T.J.; Blum, H.E.; Shafritz, D.A. Molecular pathogenesis of liver disease during persistent hepatitis B virus infection. *Semin. Liver Dis.* **1992**, *12*, 252–264. [CrossRef] [PubMed]

64. Banerjee, K.; Khandekar, P. Identification and characterization of mutations in X region of a hepatitis B virus carrier. *Indian J. Biochem. Biophys.* **1992**, *29*, 494–497. [PubMed]

65. Karayiannis, P.; Alexopoulou, A.; Hadziyannis, S.; Thursz, M.; Watts, R.; Seito, S.; Thomas, H.C. Fulminant hepatitis associated with hepatitis B virus e antigen-negative infection: Importance of host factors. *Hepatology* **1995**, *22*, 1628–1634. [PubMed]

66. Asahina, Y.; Enomoto, N.; Ogura, Y.; Kurosaki, M.; Sakuma, I.; Izumi, N.; Marumo, F.; Sato, C. Sequential changes in full-length genomes of hepatitis B virus accompanying acute exacerbation of chronic hepatitis B. *J. Hepatol.* **1996**, *25*, 787–794. [CrossRef]

67. Fang, D.G.R.; Zhang, Q.; Li, Z.; Duan, S.; Yin, Z. Transient expression and antigenic characterization of HBsAg 126Ser of hepatitis B virus aa126Ile->Ser mutant. *Chin. J. Virol.* **1998**, *14*, 1–9.

68. Bowyer, S.M.; van Staden, L.; Kew, M.C.; Sim, J.G. A unique segment of the hepatitis B virus group a genotype identified in isolates from South Africa. *J. Gen. Virol.* **1997**, *78*, 1719–1729. [CrossRef] [PubMed]

69. Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: Comparative analysis of 40 full-length isolates. *Arch. Virol.* **1998**, *143*, 2313–2326. [CrossRef] [PubMed]

70. Grethe, S.; Monazahian, M.; Bohme, I.; Thomssen, R. Characterization of unusual escape variants of hepatitis B virus isolated from a hepatitis B surface antigen-negative subject. *J. Virol.* **1998**, *72*, 7692–7696. [PubMed]

71. Mbayed, V.A.; Lopez, J.L.; Telenta, P.F.S.; Palacios, G.; Badia, I.; Ferro, A.; Galoppo, C.; Campos, R. Distribution of hepatitis B virus genotypes in two different pediatric populations from Argentina. *J. Clin. Microbiol.* **1998**, *36*, 3362–3365. [CrossRef] [PubMed]

72. Poovorawan, Y.; Sripattanawat, R.; Theamboonlers, A.; Chongsrisawat, V.; Nuchprayoon, I. Hepatocellular carcinoma: Significance of HBV vertical transmission. *Asian Pac. J. Allergy Immunol.* **1998**, *16*, 93–103. [PubMed]

73. Colonno, R.J.; Rose, R.; Baldick, C.J.; Levine, S.; Pokornowski, K.; Yu, C.F.; Walsh, A.; Fang, J.; Hsu, M.; Mazzucco, C.; et al. Entecavir resistance is rare in nucleoside naïve patients with hepatitis B. *Hepatology* **2006**, *44*, 1656–1665. [CrossRef] [PubMed]

74. Tenney, D.J.; Levine, S.M.; Rose, R.E.; Walsh, A.W.; Weinheimer, S.P.; Discotto, L.; Plym, M.; Pokornowski, K.; Yu, C.F.; Angus, P.; et al. Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to lamivudine. *Antimicrob. Agents Chemother.* **2004**, *48*, 3498–3507. [CrossRef] [PubMed]

75. Langley, D.R.; Walsh, A.W.; Baldick, C.J.; Eggers, B.J.; Rose, R.E.; Levine, S.M.; Kapur, A.J.; Colonno, R.J.; Tenney, D.J. Inhibition of hepatitis B virus polymerase by entecavir. *J. Virol.* **2007**, *81*, 3992–4001. [CrossRef] [PubMed]

76. Bartholomeusz, A.; Locarnini, S. Hepatitis B virus mutations associated with antiviral therapy. *J. Med. Virol.* **2006**, *78*, S52–S55. [CrossRef] [PubMed]

77. Baldick, C.J.; Tenney, D.J.; Mazzucco, C.E.; Eggers, B.J.; Rose, R.E.; Pokornowski, K.A.; Yu, C.F.; Colonno, R.J. Comprehensive evaluation of hepatitis B virus reverse transcriptase substitutions associated with entecavir resistance. *Hepatology* **2008**, *47*, 1473–1482. [CrossRef] [PubMed]

78. Widasari, D.I.; Yano, Y.; Heriyanto, D.S.; Utsumi, T.; Yamani, L.N.; Rinonce, H.T.; Wasityastuti, W.; Lusida, M.I.; Soetjipto; Okada, R.; et al. A deep-sequencing method detects drug-resistant mutations in the hepatitis B virus in Indonesians. *Intervirology* **2014**, *57*, 384–392. [CrossRef] [PubMed]

79. Kim, D.Y.; Chang, H.Y.; Lim, S.M.; Kim, S.U.; Park, J.Y.; Kim, J.K.; Lee, K.S.; Han, K.H.; Chon, C.Y.; Ahn, S.H. Quasispecies and pre-existing drug-resistant mutations of hepatitis B virus in patients with chronic hepatitis B. *Gut Liver* **2013**, *7*, 329–334. [CrossRef] [PubMed]

80. Ruiz-Jarabo, C.M.; Arias, A.; Baranowski, E.; Escarmis, C.; Domingo, E. Memory in viral quasispecies. *J. Virol.* **2000**, *74*, 3543–3547. [CrossRef] [PubMed]

81. Arias, A.; Ruiz-Jarabo, C.M.; Escarmis, C.; Domingo, E. Fitness increase of memory genomes in a viral quasispecies. *J. Mol. Biol.* **2004**, *339*, 405–412. [CrossRef] [PubMed]

82. Menendez-Arias, L.; Sebastian-Martin, A.; Alvarez, M. Viral reverse transcriptases. *Virus Res.* **2017**, *234*, 153–176. [CrossRef] [PubMed]

83. Ono-Nita, S.K.; Kato, N.; Shiratori, Y.; Masaki, T.; Lan, K.H.; Carrilho, F.J.; Omata, M. YMDD motif in hepatitis B virus DNA polymerase influences on replication and lamivudine resistance: A study by in vitro full-length viral DNA transfection. *Hepatology* **1999**, *29*, 939–945. [CrossRef] [PubMed]

84. Rhee, S.Y.; Margeridon-Thermet, S.; Nguyen, M.H.; Liu, T.F.; Kagan, R.M.; Beggel, B.; Verheyen, J.; Kaiser, R.; Shafer, R.W. Hepatitis B virus reverse transcriptase sequence variant database for sequence analysis and mutation discovery. *Antivir. Res.* **2010**, *88*, 269–275. [CrossRef] [PubMed]