Article

rt269I Type of Hepatitis B Virus (HBV) Polymerase versus rt269L Is More Prone to Mutations within HBV Genome in Chronic Patients Infected with Genotype C2: Evidence from Analysis of Full HBV Genotype C2 Genome

Hyein Jeong [1,†], Dong Hyun Kim [1,†], Yu-Min Choi [1], HyeLim Choi [2], Donghyun Kim [2] and Bum-Joon Kim [1,*]

1 Department of Biomedical Sciences, Microbiology and Immunology, and Liver Institute, College of Medicine, Seoul National University, Seoul 03080, Korea; zzzzihye@snu.ac.kr (H.J.); shady6233@snu.ac.kr (D.H.K.); cym486486@snu.ac.kr (Y.-M.C.)

2 Department of Biomedical Sciences, and Microbiology and Immunology, College of Medicine, Seoul National University, Seoul 03080, Korea; helenchoi501@gmail.com (H.C.); biologokim@snu.ac.kr (D.K.)

* Correspondence: kbumjoon@snu.ac.kr

† These authors have contributed equally to this work and share first authorship.

Abstract: Recently, it has been reported that the rt269I type of hepatitis B virus (HBV) polymerase (Pol) versus the rt269L type is more significantly related to lower viral replication and HBeAg negative infections in chronic hepatitis B (CHB) patients of genotype C2. In this study, we compared mutation rates within HBV genomes between rt269L and rt269I using a total of 234 HBV genotype C2 full genome sequences randomly selected from the HBV database (115 of rt269L and 119 of rt269I type). When we applied the Benjamini and Hochberg procedure for multiple comparisons, two parameters, dN and d, at the amino acids level in the Pol region were significantly higher in the rt269I type than in the rt269L type. Although it could not reach statistical significance from the Benjamini and Hochberg procedure, nonsynonymous (NS) mutations in the major hydrophilic region (MHR) or "a" determinant in the surface antigens (HBsAg ORF) related to host immune escape or vaccine escape are more frequently generated in rt269I strains than in rt269L. We also found that there are a total of 19 signature single nucleotide polymorphisms (SNPs), of which 2 and 17 nonsynonymous mutation types were specific to rt269L and rt269I, respectively: Of these, most are HBeAg negative infections (preC-W28*, X-V5M and V131I), lowered HBV DNA or virion production (C-I97F/L, rtM204I/V) or preexisting nucleot(s)ide analog resistance (NAr) (rtN139K/H, rtM204I/V and rtI224V) or disease severity (preC-W28*, C-I97F/L, C-Q182K/*, preS2-F141L, S-L213I/S, V/L5M, T36P/S/A, V131I, rtN139K/H, rtM204I/V and rtI224V). In conclusion, our data showed that rt269I types versus rt269L types are more prone to overall genome mutations, particularly in the Pol region and in the MHR or "a" determinant in genotype C2 infections and are more prevalent in signature NS mutations related to lowered HBV DNA replication, HBsAg and HBeAg secretion and potential NAr variants and hepatocellular carcinoma (HCC), possibly via type I interferon (IFN-I)-mediated enhanced inflammation. Our data suggest that rt269L types could contribute to liver disease progression via the generation of immune escape or enhanced persistent infection in chronic patients of genotype C2.

Keywords: rt269L-type hepatitis B virus (HBV); rt269I; nonsynonymous (NS) mutation; genotype C2

1. Introduction

Despite the availability of an effective vaccine, hepatitis B virus (HBV) infection is still a high-risk global health issue; with more than 240 million people being chronic carriers of the virus, approximately 786,000 patients annually die worldwide due to hepatitis B virus (HBV)-related diseases, including cirrhosis and hepatocellular carcinoma (HCC) [1]. HBV infection is endemic in South Korea, with the prevalence of hepatitis B virus surface

antigen (HBsAg) positivity being 3.4% among men and 2.6% among women on the basis of the Korean National Health and Nutrition Survey of 2011 [2].

HBV has an enveloped and partially double-stranded DNA virus and preferentially replicates in hepatocytes with an incomplete double-stranded DNA genome of approximately 3.2 kb in length; it contains four overlapping open reading frames (ORFs): surface antigens (S), core proteins (C), polymerase (Pol), and X proteins (X) [3]. HBV reverse transcriptase (RT) lacks proofreading ability, leading to HBV mutations that occur at a 10-fold higher frequency than other DNA viruses [4]. This compromises antiviral therapy with nucleos(t)ide analogs (NAs) and affects disease progression via persistent infections.

On the basis of an 8% divergence in HBV genome sequences, HBV has been characterized into 10 genotypes as A–J [5]. There is increasing evidence that specific HBV genotypes may play significant roles in the development of different disease profiles during chronic hepatitis B (CH) infection [6,7]. Among the 10 genotypes, genotypes B and C are widespread in Asia, but two genotypes lead to distinctly different clinical outcomes [8]. Compared to genotype B, genotype C, particularly C2, showed higher HBV replication capacity, a higher tendency of chronicity and more frequently developed into liver cirrhosis (LC) and hepatocellular carcinoma (HCC) [9]. In addition, incomplete response to interferon (IFN) therapy and higher levels of mutations were also reported in genotype C infection. Notably, an extraordinary prevalence of virulent genotype C2 has been reported in South Korea [10–12]. Furthermore, the presence of a distinct immune response against HBV proteins in the Korean population can lead to the generation of unique HBV variants that are rarely encountered in other areas, resulting in distinct clinical manifestations in Korean chronic patients [13]. Indeed, several unique types of HBV mutations related to the progression of liver disease encountered in other areas have been found in South Korea [14–29]. To date, higher mutation rates and the presence of distinct mutations found in genotype C2 infections remain largely unknown. However, recently, we have reported that the presence of two HBV Pol RT polymorphisms, rt269L and rt269I, distinct only in HBV genotype C, could play a very pivotal role in viral phenotypes, clinical outcomes, and worse responses to IFN therapy distinct in genotype C2 infections. In particular, we showed that rt269I-type infection versus rt269L-type infection leads to enhanced mitochondrial stress-mediated type I interferon (IFN-I) production, resulting in HBV e antigen (HBeAg)-negative infections by generating preC mutations at 1896 (G to A). Therefore, in this study, we hypothesized that HBeAg seroconversion, frequently found in patients infected with the rt269I type of genotype C2, could lead to increased genome mutations in the rt269I type versus the rt269L type by inducing immune responses against HBV. To address this issue, we compared overall mutation rates and signature mutation SNPs specific to the respective RT polymorphisms rt269L and rt269I using a total of 234 HBV genotype C2 full genome sequences retrieved from a public database (Hepatitis B Virus Database: https://hbvdb.ibcp.fr/HBVdb/ (accessed on 14 March 2021)), of which 115 and 119 genome sequences were from rt269L and rt269I types, respectively.

2. Materials and Methods

2.1. HBV Sequence Acquisition and Processing

First, a total of 252 HBV genotype C2 sequence data were downloaded from the Hepatitis B Virus Database (https://hbvdb.ibcp.fr/HBVdb/ (accessed on 14 March 2021)) [30]. The accession numbers of downloaded sequences are shown in Supplementary Table S1. Of the 252 HBV sequences, 234 except for 18 strains identified as subgenotypes, C3 and C4, by our full genome-based phylogenetic analysis were finally determined as genuine genotype C2 strains. For the analysis of mutation rates, each of the five genes (preS, S, C, P, and X) was extracted and aligned separately using the MEGA X program (USA). Every gap and insert were deleted and aligned, so the whole length of the sequences was fit to 3215 bp. All the sequence sites were numbered for the *EcoR*1 restriction site to be located as a first nucleotide (nt). For the purpose of the study, every sequence was checked at the 269th codon on reverse transcriptase to classify the L type (rt269L) and I type (rt269I).

2.2. HBV Phylogenetic Analysis and Intergenotypic Recombination Analysis

A phylogenetic analysis based on entire sequences of HBV full genome sequences (3215 bp) was performed for further subgenotype separation of 252 HBV designated genotype C2 from HBVdb [30]. The analysis was conducted with five genotype reference strains (M57663 (A), AB100695 (B), X02496 (D), AB106564 (E), and X75663 (F)) and thirteen recently described subgenotype reference strains [31] (including C1 (AB074047, EU306686, GQ924619), C2 (KM99991), C3 (AB115417, AB115418, AY167091), C4 (EU939624, AY206374, AB675675), and C5 (GQ924620, GQ924657, AP011100) obtained from GenBank. A phylogenetic tree was inferred by the neighbor-joining method using Tamura–Nei genetic distance model and 100 replicates of bootstrap tests were conducted [32–34]. HBV genotype B sequences were used as the outgroup (accession No. D00329). The intergenotypic recombination of sequences of the 234 strains finally identified as genuine subgenotype C2 were analyzed with five genotype reference strains (M57663 (A), AB100695 (B), X02496 (D), AB106564 (E), and X75663 (F)) using DualBrothers recombination detection software [35].

2.3. Analysis of Mutation Rates Based on the HBV Full Genome and 5 HBV Regions (preS1, S, preC/C, X and Pol)

Genetic diversity, including the mean genetic distance (d), the number of synonymous substitutions per synonymous site (dS), and the number of nonsynonymous substitutions per nonsynonymous site (dN), was calculated for each reference strain of 234 HBV sequences finally determined as genotype C2 infections. The mean genetic distance was a measure of the accumulated number of nucleotide or codon differences per locus from the consensus sequence. The consensus sequence was defined as the calculated order of most frequent residues, either nucleotide or amino acid, found at each position in a sequence alignment. The consensus sequence was differently calculated in 115 rt269L and 119 rt269I strains, but if the consensus sites obtained in rt269L and rt269I were different at a nucleotide or amino acid position, the consensus at the position included both of them. Exceptionally, adenine substitution at 1762 nt and guanine substitution at 1764 nt, which have been widely accepted as wild-type sequences in the basal core promoter (BCP) [36], were defined as wild type. The consensus sequences were calculated in each open reading frame (PreS, S, preC/C, X, and Pol). The number of amino acid variants was calculated as the sum of those in each ORF, taking into account the partially overlapping ORFs of HBV.

The genetic distances at the nucleotide level were calculated under the Tamura three-parameter model, taking into account the differences in transitional and transversional rates and G+C-content bias. Meanwhile, the genetic distances at the amino acid level were calculated under the Jones–Taylor–Thorton (JTT) model. The number of both nucleotide and amino acid variants was calculated. The dS and dN were calculated under the modified Nei–Gojobori model with Jukes–Cantor correction. All genetic diversity factors were tested using MEGA X software (USA) [37].

2.4. Statistical Analysis

The results of continuous variables are expressed as medians and standard deviations. The genetic diversity between rt269L and rt269I was compared using independent sample *t*-tests. Mutation rate of signature nonsynonymous mutation is defined as the proportion of the number of strains among each group. The signature of nonsynonymous mutation between rt269L and rt269I was compared using chi-squared test. The statistical analysis was performed using GraphPad Prism 8.0 software (GraphPad, La Jolla, CA, USA). For the analysis of signature nonsynonymous mutations, continuous variables were tested using the chi-squared test using SPSS 25.0.0.0 software (SPSS Inc., Chicago, IL, USA). Significant differences (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) among the different groups are shown in the related figures, and the bar plot data are presented as the means $\pm$ s.e.m. of each group. Additionally, multiple comparison using Benjamini–Hochberg correction (false discovery rate (FDR)) was also performed by R 3.5.2 software and the significant values ([a] $q < 0.05$, [b] $q < 0.01$, [c] $q < 0.001$) are shown in the related tables [38,39].

3. Results

3.1. Genotype Determination of HBV Full Genome Sequences via Phylogenetic Analysis

To analyze overall genome mutation rates and signature SNPs between HBV genotype C2 strains of two RT polymorphisms, rt269L and rt269I, 252 HBV strains belonging to genotype C2 were randomly selected from the HBV database HBVdb (http://hbvdb.ibcp.fr (accessed on 14 March 2021)) [30]. Since 18 (2 strains of rt269L and 16 strains of rt269I) of 252 strains designated genotype C2 in HBVdb were identified into subgenotypes, C3 and C4, by our phylogenetic analysis based on a recent study newly defining HBV genotypes and subgenotypes [31,40,41] (data not shown), we finally determined 234 HBV strains (92.9%) as genotype C2, which were used for our genome study. They were further separated into 115 rt269L types (49.1%) encoding leucine and 119 rt269I types (50.9%) encoding isoleucine at the 269th codon of RT via sequence comparison (Figure 1 and Supplementary Table S1).
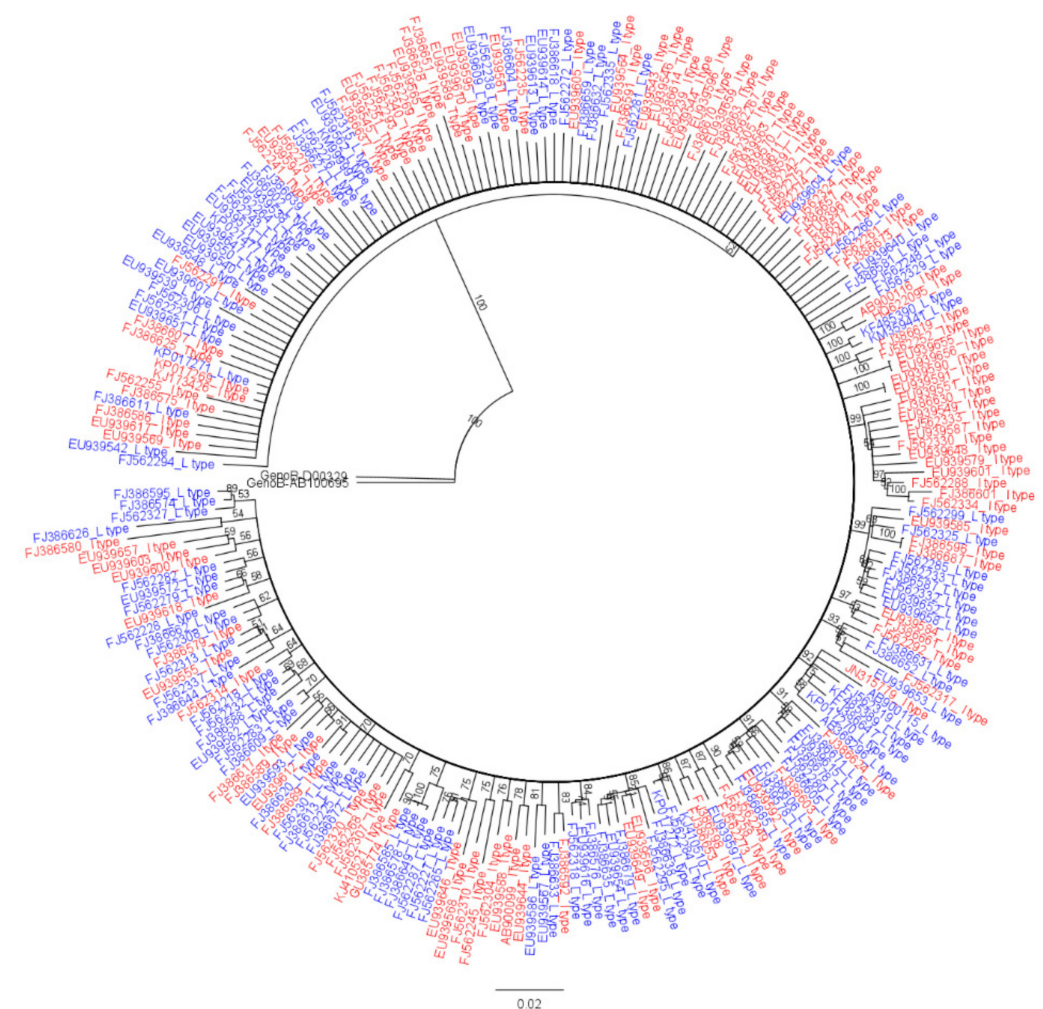

Figure 1. Genotype determination of selected HBV genomes via phylogenetic analysis. Phylogenetic tree of a total of 234 HBV sequences belonging to genotype C2 was conducted by the neighbor-joining method using Tamura–Nei genetic distance model. A total of 100 replicate bootstrap tests were conducted using a genotype B sequence as an outgroup. rt269L and rt269I sequences labeled "L" and "I" at the end of accession number.

3.2. Comparison of Genetic Diversity of the HBV Full Genome between rt269L and rt269I

The genome diversity of the HBV genome between rt269L and rt269I, including the mean genetic distance (d), the number of synonymous substitutions per synonymous site (dS), and the number of nonsynonymous substitutions per nonsynonymous site (dN), was calculated for each HBV strain using independent sample *t*-test. As HBV partially

overlaps four open reading frames encoding the surface antigen, nucleocapsid, X protein (X) and polymerase (Pol), we separately estimated the total number (1613 aa) of amino acid mutations in the presurface (PreS) domain, S protein, X, precore/core (PreC/C), and Pol coding regions of the HBV genome. Our statistical analysis using *t*-test showed that the numbers of nucleotide and amino acid variants in the HBV full genome of the rt269L type (38.03 $\pm$ 13.65/3215 bp and 23.43 $\pm$ 10.75 variants/1613 aa, respectively) were significantly lower than those of the rt269I type (42.34 $\pm$ 15.17/3215 bp and 27.68 $\pm$ 12.39/1613 aa variants, respectively) (* *p*= 0.0234 and ** *p* = 0.0056, respectively) (Table 1 and Figure 2).

Table 1. Comparison of genetic diversity within the hepatitis B virus (HBV) full genome, PreS, S, X, PreC/C and Pol regions between rt269L and rt269I.

	L Type (*n* = 115)	I Type (*n* = 119)	*p*-Value from *t*-Test
Number of nucleotide variants			
HBV full genome (/3,215 bp)	38.03 $\pm$ 13.65	42.34 $\pm$ 15.17	0.0234 *
Number of amino acid variants			
HBV full genome (/1,613 aa)	23.43 $\pm$ 10.75	27.68 $\pm$ 12.39	0.0056 **
d (10^{-2} substitution/site) (nucleotide level)			
preS	1.39 $\pm$ 1.59	1.71 $\pm$ 1.78	0.1559
SHB	0.58 $\pm$ 0.37	0.64 $\pm$ 0.45	0.2522
pol	1.09 $\pm$ 0.46	1.24 $\pm$ 0.50	0.0154 *
RT	0.97 $\pm$ 0.40	1.03 $\pm$ 0.48	0.3216
X	1.24 $\pm$ 0.61	1.28 $\pm$ 0.60	0.6752
preC/C	1.38 $\pm$ 0.86	1.56 $\pm$ 0.80	0.0989
d (10^{-2} substitution/site) (amino acid level)			
preS	1.60 $\pm$ 2.50	2.23 $\pm$ 3.29	0.0989
SHB	1.10 $\pm$ 0.95	1.29 $\pm$ 1.01	0.1494
pol	1.26 $\pm$ 0.79	1.60 $\pm$ 0.86	0.0019 **[a]
RT	0.88 $\pm$ 0.67	1.06 $\pm$ 0.72	0.0465 *
X	2.19 $\pm$ 1.43	2.34 $\pm$ 1.30	0.3796
preC/C	1.64 $\pm$ 1.56	1.66 $\pm$ 1.29	0.9128
dN (10^{-2} substitution/site)			
preS	0.81 $\pm$ 1.40	1.11 $\pm$ 1.74	0.1449
SHB	0.52 $\pm$ 0.45	0.62 $\pm$ 0.49	0.1142
pol	0.61 $\pm$ 0.42	0.78 $\pm$ 0.45	0.0044 **[a]
RT	0.42 $\pm$ 0.32	0.51 $\pm$ 0.35	0.0327 *
X	1.04 $\pm$ 0.70	1.10 $\pm$ 0.62	0.4405
preC/C	0.82 $\pm$ 0.81	0.82 $\pm$ 0.65	0.9903
dS (10^{-2} substitution/site)			
preS	2.78 $\pm$ 2.23	3.22 $\pm$ 2.07	0.1188
SHB	0.68 $\pm$ 0.55	0.69 $\pm$ 0.65	0.9058
pol	2.21 $\pm$ 0.74	2.31 $\pm$ 0.86	0.2855
RT	2.30 $\pm$ 1.05	2.26 $\pm$ 1.23	0.8048
X	1.72 $\pm$ 1.05	1.68 $\pm$ 0.90	0.7902
preC/C	2.45 $\pm$ 1.62	2.94 $\pm$ 1.62	0.0212 *

d, genetic distance; dS, the number of synonymous substitutions per synonymous site; dN, the number of non-synonymous substitutions per non-synonymous site; the significant values from *t*-test were shown in boldface and marked with asterisk (* *p* < 0.05, ** *p* < 0.01); [a] Statistically significant after Benjamini–Hochberg false discovery rate (FDR) ([a] *q* < 0.05).
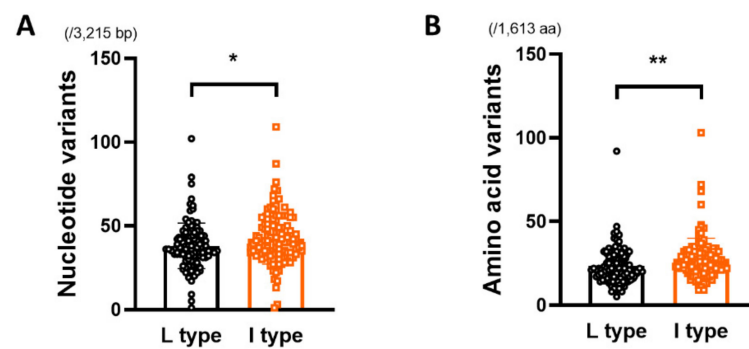
Figure 2. Comparison of genetic diversity of the HBV full genome between rt269L and rt269I. Comparison of the number of (**A**) nucleotide and (**B**) amino acid variants within the HBV full genome between rt269L and rt269I. Significant differences (* $p < 0.05$, ** $p < 0.01$) analyzed by independent samples *t*-test between the groups are shown in the related figures. The bar plot data are presented as the mean $\pm$ s.e.m. of each group.

3.3. Comparison of Genetic Diversity in the PreS, S, X, PreC/C and Pol Regions between rt269L and rt269I

The genetic diversity within the PreS, S, X, PreC/C, and Pol regions, including the nucleotide and amino acid levels, was calculated for each strain in both rt269L and rt269I using the independent sample *t*-test. Our statistical analysis using the *t*-test showed that in three regions—the PreS, S, and X regions—there were no significant differences between rt269L and rt269I types in d at either the nucleotide or amino acid level, dN, and dS ($p > 0.05$). However, in the Pol region, dNS and d at both the nucleotide and amino acid levels were significantly lower in the rt269L type (0.61 ± 0.42, 1.09 ± 0.46, and 1.26 ± 0.79 $\times 10^{-2}$ substitutions/site, respectively) than in the rt269I type (0.78 ± 0.45, 1.24 ± 0.50, and 1.60 ± 0.86 $\times 10^{-2}$ substitutions/site, respectively) (** $p = 0.0044$, * $p = 0.0154$, and ** $p = 0.0019$ from *t*-test, respectively). In particular, in the RT region of HBV polymerase, dN and d at the amino acid level were significantly lower in the rt269L type (0.42 ± 0.32 and 0.88 ± 0.67 $\times 10^{-2}$ substitutions/site, respectively) than in the rt269I type (0.51 ± 0.35 and 1.06 ± 0.72 $\times 10^{-2}$ substitutions/site, respectively) (* $p = 0.0327$ and * $p = 0.0465$ from *t*-test, respectively). In the PreC/C region, dS and d at the nucleotide level were significantly or tended to be lower in the rt269L type (2.45 ± 1.62 and 1.38 ± 0.86 $\times 10^{-2}$ substitutions/site, respectively) than in the rt269I type (2.94 ± 1.62 and 1.56 ± 0.80 $\times 10^{-2}$ substitutions/site, respectively) (* $p = 0.0212$ and $p = 0.0989$ from *t*-test, respectively) but not different in dNS and d at the amino acid level (Table 1 and Figure 3). Moreover, when we applied the Benjamini and Hochberg procedure to adjust the raw *p*-values for multiple comparisons, notably, two parameters, dN and d, at the amino acids level in the Pol region still had significant difference, although most corrected *p*-values did not reach the statistical significance (Table 1). Together, our data indicated that rt269I types versus rt269L are more prone to overall genome mutations, particularly in the Pol region.
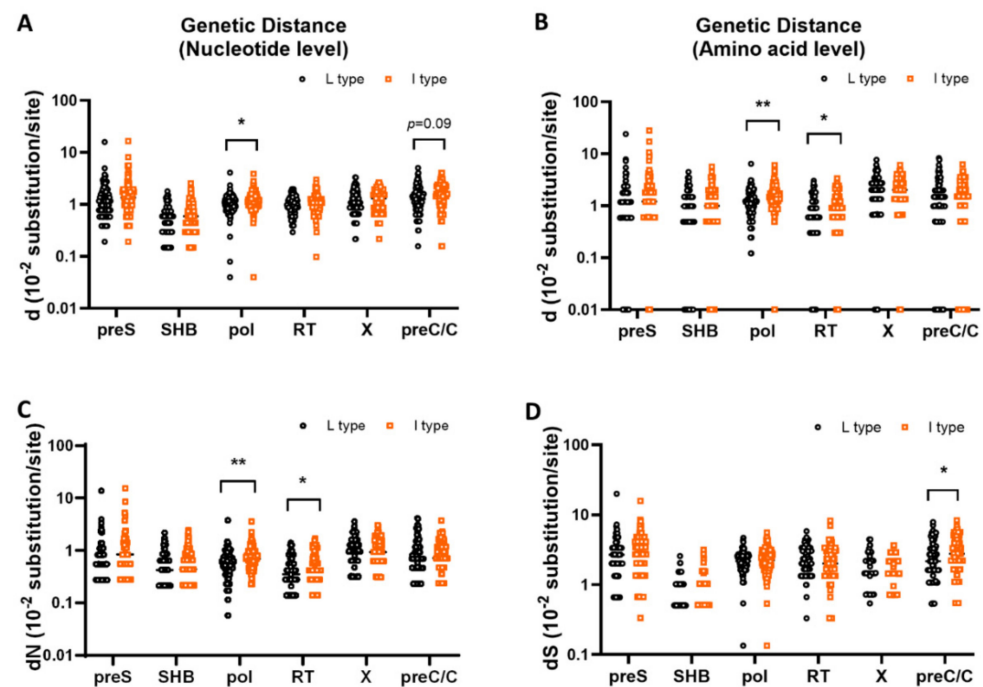
Figure 3. Comparison of genetic diversity in the PreS, S, X, PreC/C and Pol regions between rt269L and rt269I. Comparison of genetic distance at the nucleotide and amino acid levels (**A**,**B**), dN (the number of nonsynonymous substitutions per nonsynonymous site) (**C**), and dS (the number of synonymous substitutions per synonymous site) (**D**) within the PreS, S, X, PreC/C and Pol regions between rt269L and rt269I. Significant differences (* $p < 0.05$, ** $p < 0.01$) analyzed by independent samples *t*-test between the groups are shown in the related figures.

3.4. rt269I versus rt269L Is More Prone to NS Mutations in the MHR or "a" Determinant Region

The major hydrophilic region (MHR) ranges from residues 99 to 169 of HBsAg, and because it is probably composed of several B cell epitopes, mutations within the MHR, especially in the "a" determinant region (residues 124–147), can affect the antigenicity of HBsAg, resulting in the escape of the virus from neutralizing antibody responses and persistent viral infections [42]. The prevalence and pattern of MHR mutations between rt269L and rt269I strains were investigated (Table 2 and Figure 4). Our statistical analysis using *t*-test showed that the number of nucleotides in the "a" determinant and the amino acid variants in the MHR between rt269L and rt269I were significantly higher in rt269I (0.29 ± 0.59 and $0.50 \pm 0.92 \times 10^{-2}$ variants in the "a" determinant, 0.24 ± 0.54 and $0.41 \pm 0.69 \times 10^{-2}$ variants in the MHR, respectively) than in rt269L (* $p = 0.0311$ and * $p = 0.0399$ from *t*-test, respectively). Further analysis showed that dS was not significantly different. dN in both MHR and "a" determinant were significantly lower in rt269L type (0.07 ± 0.38 and $0.38 \pm 1.24 \times 10^{-2}$ substitutions/site in "a" determinant, 0.48 ± 0.53 and $0.67 \pm 0.87 \times 10^{-2}$ substitutions/site in MHR, respectively) than in rt269I type (* $p = 0.0109$ and $p = 0.0497$ from *t*-test, respectively). However, all these differences between rt269L and rt269I shown by *t*-test did not reach statistical significance in multiple comparisons using Benjamini–Hochberg correction. Together, while it may not be so obvious, NS mutations in the MHR or "a" determinant in the S ORF related to host immune escape or vaccine escape are more frequently generated in rt269I strains than in rt269L, suggesting that the former is more likely to lead to persistent infection in chronic patients with genotype C2 than the latter.

Table 2. Comparison of genetic diversity within MHR and "a" determinant regions between rt269L and rt269I.

	L Type (*n* = 115)	I Type (*n* = 119)	*p*-Value from *t*-Test
Number of nucleotide variants			
a determinant (/72 bp)	0.29 ± 0.59	0.50 ± 0.92	0.0311 *
MHR (/213 bp)	1.14 ± 1.10	1.31 ± 1.38	0.2935
Number of amino acid variants			
a determinant (/24 aa)	0.23 ± 0.45	0.33 ± 0.68	0.2182
MHR (/71 aa)	0.24 ± 0.54	0.41 ± 0.69	0.0399 *
dN (10^{-2} substitution/site)			
a determinant	0.07 ± 0.38	0.38 ± 1.24	0.0109 *
MHR	0.48 ± 0.53	0.67 ± 0.87	0.0497 *
dS (10^{-2} substitution/site)			
a determinant	0.99 ± 2.58	1.14 ± 2.36	0.6425
MHR	1.39 ± 1.48	1.26 ± 1.45	0.5065

Abbreviations: d, genetic distance; dS, the number of synonymous substitutions per synonymous site; dN, the number of non-synonymous substitutions per non-synonymous site; the significant values from *t*-test were shown in boldface and marked with asterisk (* *p* < 0.05).

3.5. Identification of Signature Nonsynonymous Mutations Specific to rt269L or rt269I

Next, we identified signature nonsynonymous mutations specific to rt269L or rt269I in the PreS, S, X, PreC/C and Pol regions, of which frequencies are significant or tend to be higher in rt269L or rt269I. The mutation rate of each signature nonsynonymous mutation was defined as the proportion of the number of strains among each group, and mutation rate between rt269L and rt269I was compared using chi-squared test. Thereafter, the resulting *p*-values were calculated by multiple comparison using Benjamini–Hochberg correction (FDR). A total of 19 types of signature nonsynonymous mutations were found in five HBV regions, PreS (three types), S (two types), X (four types), PreC/C (three types) and Pol regions (seven types). Of these, 2 and 17 nonsynonymous mutation types were specific to rt269L and rt269I, respectively (Table 3).

3.5.1. PreS and S Region

HBV S ORF is composed of preS1, preS2 and S regions. NS mutations in these regions could lead to liver disease progression by escaping host immune responses and eliciting endoplasmic reticulum stress [43]. We found a total of five types of signature NS mutations in the S ORF (I84T/M/L/V and A90T/V in PreS1, F141L in PreS2, and I68T and L213I/S in S). Of these, the frequency in all four types but one rt269L signature type (I68T in the S region) was significantly or tended to be higher in rt269I HBV strains than in rt269L. Of the four rt269I signature types, I84T/M/L/V in PreS1 has been reported to be frequently found in HBV occult cases [15]. F141L in PreS2 has been reported to be significantly related to HCC progression in Korean chronic patients by inducing cell cycle progression by downregulating the p53 and p21 pathways and upregulating CDK4 and cyclin A [28]. The L213I/S in the S region, which tended to be higher in rt269I HBV strains than in rt269L (*q* = 0.054), has been reported to be related to disease progression in chronic patients) [44]. The rt269L signature type, I68T in the S region, which is significantly prevalent in the rt269L versus rt269I type (*q* = 0.004), has been reported to be an HCC-independent risk factor in chronic Chinese patients with genotype C infections [45].
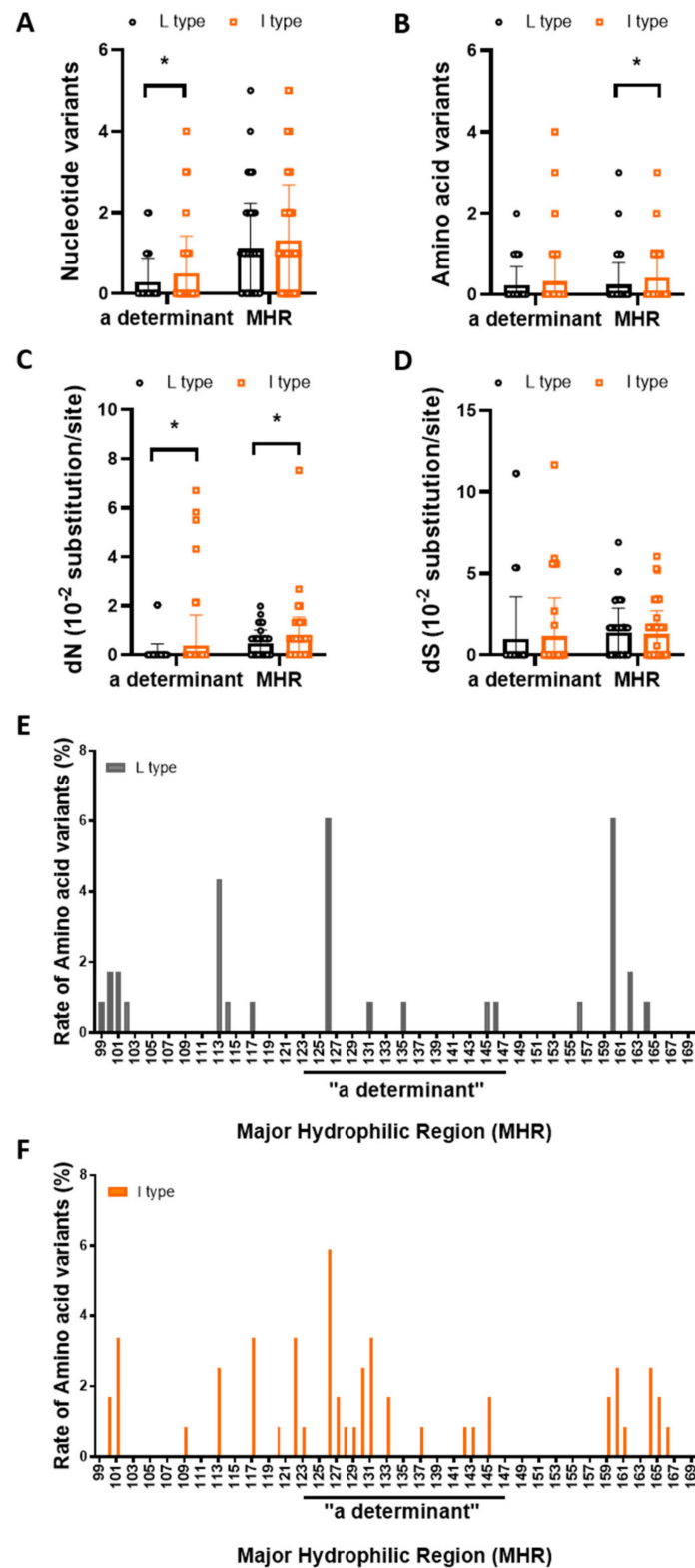
Figure 4. rt269I versus rt269L is more prone to NS mutations in the MHR or "a" determinant region. Comparison of the number of nucleotide and amino acid variants (**A**,**B**), dN (**C**), and dS (**D**) within the major hydrophilic region and the "a" determinant. Significant differences (* $p < 0.05$) analyzed by independent samples *t*-test between the groups are shown in the related figures. Prevalence of residue substitutions within MHR from rt269L (black bars) and rt269I (orange bars) (**E**,**F**). Each bar represents the rate of amino acid variants located on a given amino acid position.

3.5.2. X Region

Since the HBV X region includes not only the X protein coding region but also the regulatory region, mutation in the X region can simultaneously affect both the transactivating activity of HBxAg and the regulation of HBV replication [46]. We found a total of four types of signature NS mutations in the X ORF (V/L5M, T36P/S/A, S101P/A and V131I). Of these, three NS mutations, V/L5M, S101P/A and V131I, and one, T36P/S/A, were rt269I and rt269L signature types, respectively. V/L5M, which tended to be higher in rt269I strains than in rt269L strains, has been reported to be significantly related to HCC progression and HBeAg-negative serostatus in chronic patients with genotype C2 infections) [24]. V131I also leads to one of the double mutations in the overlapping basal core promoter (1764 G to A), which is significantly prevalent in rt269I strains versus rt269L ($q = 0.004$) and has also been reported to be related to HBsAg-negative serostatus and HCC progression in chronic patients [47]. T36P/S/A, a rt269L signature type that is more prevalent in rt269I strains than in rt269L strains ($q < 0.001$), has been reported to be higher in HCC patients, which enables the escape of the host immune response, possibly due to its location at the B epitope [48].

3.5.3. PreC/C region

Mutations in the preC/C region can affect HBeAg serostatus and antigenicity, HBV nucleocapsid structure and stability, and the packaging of pregenomic RNA into the nucleocapsid [49]. Furthermore, HBcAg is the principal target of the host cell-mediated immune response [50,51] of which mutations can induce persistent HBV infections) [14,52]. We found a total of three types of signature NS mutations in the preC/C region (one in the preC region: preC-W28Stop, two in the C region: C-I97F/L (A2189T/C or C2191T mutation) and C-Q182K/* (C2444A/T or A2445G mutation)). All the three were rt269I signature types. The preC-W28Stop, known as a preC mutation at 1896 nt, leads to premature termination of HBeAg, resulting in HBeAg-negative infection in chronic patients with its mutation, of which the frequency was significantly higher in rt269I strains than in rt269L strains ($q < 0.001$). A number of studies have reported that preC-W28Stop, a preC 1896 mutation, is significantly related to liver disease progression [53,54]. The C-I97F/L in the C region, of which the frequency was significantly higher in rt269I strains than in rt269L strains ($q = 0.047$), is well known as the most frequently encountered HBcAg mutation, as mentioned in several studies [55,56] leading to an immature secretion phenotype due to defective nucleocapsid assembly. This can contribute to the progression of severe liver disease by failing to elicit a proper host immune response against HBV infection. Indeed, a previous study reported that C-I97F/L (A2189T/C or C2191T mutation) was the preC/C type found the most frequently in HCC patients in Korean chronic patients infected with genotype C2 [13,14]. C-Q182K/* (C2444A/T or A2445G mutation), of which the frequency tended to be higher in rt269I strains than in rt269L strains ($q = 0.068$), has been reported to be significantly related to HCC in Korean chronic patients infected with genotype C2 [13,14].

Table 3. Comparison of the signature nonsynonymous mutation rate between rt269L and rt269I.

Mutation Region	Amino Acid	Reference Strains		FDR Corrected p-Value (q-Value)	Function	Ref
		L Type (*n* = 115)	I Type (*n* = 119)			
preS1	I84T/M/L/V	Ile (I) 99 (86.09%)	82 (68.91%)	<0.001 [c]	Transactivator domain, Occult infection	[17]
		Thr (T) 9 (7.83%)	18 (15.13%)			
		Met (M) 1 (0.87%)	3 (2.52%)			
		Leu (L) 1 (0.87%)	14 (11.76%)			
		Val (V) 2 (1.74%)	2 (1.68%)			
	A90T/V	Ala (A) 84 (73%)	57 (47.9%)	<0.001 [c]		
		Th r(T)	1 (0.8%)			
		Val (V) 31 (27%)	61 (51.3%)			
preS2	F141L	Phe (F) 100 (86.96%)	95 (79.83%)	0.046 [a]	HCC	[27]
		Leu (L) 12 (10.43%)	24 (20.17%)			
S	sI68T	Ile (I) 81 (70.43%)	103 (86.55%)	0.004 [b]	HCC	[57]
		Thr (T) 34 (29.57%)	16 (13.45%)			
	sL213I/S	Leu (L) 110 (96.65%)	105 (88.24%)	0.054	Disease progression	[58]
		Ile (I) 5 (4.35%)	13 (10.92%)			
		Ser (S) 0 (0.00%)	1 (0.84%)			
Terminal protein	D16A/E/N/V	Asp (D) 110 (95.65%)	105 (88.24%)	0.054	N-terminal helix1 mutation	[59]
		Ala (A) 1 (0.87%)	3 (2.52%)			
		Glu (E) 2 (1.74%)	4 (3.36%)			
		Asn (N) 1 (0.87%)	3 (2.52%)			
		Val (V) 1 (0.87%)	4 (3.36%)			
Spacer	S314P	Ser (S) 78 (67.83%)	50 (42.01%)	<0.001 [c]		
		Pro (P) 37 (32.17%)	69 (57.98%)			
	F321L	Phe (F) 79 (68.70%)	51 (42.86%)	<0.001 [c]		
		Leu (L) 36 (31.30%)	68 (57.14%)			
Reverse transcriptase	rtN139K/H	Asn (N) 112 (98.26%)	112 (94.96%)	0.066	Pretreatment mutation	[60]
		Lys (K) 0(0.00%)	2 (1.68%)			
		His (H) 1 (1.74%)	5 (4.20%)			
	rtM204I/V	Met(M) 111(96.52%)	103 (86.55%)	0.009 [b]	resistance to LMV, LdT, ADV, TNF	[61,62]
		Ile (I) 4 (3.48%)	14 (11.76%)			
		Val (V) 0 (0.00%)	2 (1.68%)			
	rtI224V	Ile (I) 110 (95.65%)	106 (89.08%)	0.042 [a]	Pretreatment mutation	[60]
		Val (V) 4 (3.48%)	13 (10.92%)			

Table 3. *Cont.*

Mutation Region	Amino Acid	Reference Strains		FDR Corrected p-Value (q-Value)	Function	Ref
		L Type (*n* = 115)	I Type (*n* = 119)			
RNaseH	D828A/V	Asp (D) 105(91.30%)	88 (73.95%)	0.001 [b]	putative catalytic site mutants	[63]
		Ala (A) 6(5.22%)	5 (4.20%)			
		Val (V) 4 (3.48%)	26 (21.85%)			
HBx	V/L5M	Val (V) 98 (85.22%)	92 (77.31%)	0.054	Independent predictors of HCC survival, miRNA binding site	[11]
		Leu (L) 6 (5.22%)	6 (5.04%)			
		Met (M) 10 (8.7%)	21 (17.65%)			
	T36P/S/A	Thr (T) 57 (50%)	89 (75.42%)	<0.001 [c]	B cell epitope	[64]
		Pro (P) 22 (19.3%)	7 (5.93%)			
		Ser (S) 22 (19.3%)	9 (7.63%)			
		Ala (A) 13 (11.4%)	13 (11.02%)			
HBx	S101P/A	Ser (S) 103(89.6%)	78 (65.5%)	<0.001 [c]		
		Ala (A) 10 (8.7%)	18 (15.1%)			
		Pro (P) 1 (0.90%)	23 (19.3%)			
	V131I	Val (V) 30 (26.79%)	14 (11.76%)	0.004 [b]	HCC	[65]
		Ile (I) 82(73.21%)	105 (88.24%)			
Precore	G1896A	Trp (W) 84 (73.04%)	61 (51.26%)	<0.001 [c]	HBeAg-negative serostatus	[14]
		Stop 31 (26.96%)	58 (48.74%)			
Core	A2189T/C, C2191T	Ile (I) 75 (65.22%)	62 (52.10%)	0.047 [a]	HBeAg-negative serostatus, HCC-related HBcAg mutation	[14,66]
		Phe (F) 5 (4.35%)	4 (3.36%)			
		Leu (L) 35 (30.43%)	53 (44.54%)			
	C2444A/T, A2445G	Gln (Q) 111 (96.52%)	107 (89.92%)	0.068	HCC-related HBcAg mutation	[14]
		Tyr (Y) 1 (0.87%)	2 (1.68%)			
		His (H) 0 (0.00%)	4 (3.36%)			
		Stop 3 (2.61%)	6 (5.04%)			

[a,b,c] Statistically significant after Benjamini–Hochberg FDR ([a] $q < 0.05$, [b] $q < 0.01$, [c] $q < 0.001$).

3.5.4. Pol Region

Preexisting HBV Pol mutations in treatment-naïve patients are related to potential drug resistance, progression of liver disease, such as HCC or cirrhosis, and final clinical outcomes of drug treatment [57]. Furthermore, several studies reported that patients with chronic HBV carrying preexisting Pol mutations had significantly decreased serum baseline HBV DNA loads and HBeAg seronegative infections [67,68]. We found a total of seven types of signature NS mutations in the Pol region (three types in RT, one type in the terminal protein (TP) region, two types in the spacer region and one type in the RNaseH region), of which all seven (rtN139K/H, rtM204I/V and rtI224V in RT, D16A/E/N/V in TP, S314P and F321L in Spacer and D828A/V in the RNaseH region) were rt269I signature types. Three NS mutation types in the RT region have been reported to belong to preexisting reverse transcriptase (RT) mutations related to nucleos(t)ide analog (NAr) resistance. rtM204I, a YMDD motif mutation whose frequency was significantly higher in rt269I than in rt269L ($q = 0.009$), can lead to resistance to various drugs, including lamivudine (LMV), and it is the most frequently encountered preexisting NAr mutation worldwide [62,69]. Of note, of the seven types, the following three rt269I signature types, rtN139K/H, rtM204I/V and rtI224V, have been reported to be NAr RT mutation types related to clinical severity, including HCC or liver cirrhosis [68,70–73]

4. Discussion

HBV genotype C2 has two polymorphisms, L and I, at the 269th codon of Pol RT, distinct from other genotypes having one I type alone, that can provide in part a likely explanation into unique virological and clinical traits found in genotype C2 infections such as higher virulence, higher mutation frequency, or lower response to IFN-I therapy [74,75] Indeed, we have previously reported that the rt269I type of genotype C2, but not genotype A, can lead to enhanced IFN-I production in infected hepatocytes, resulting in HBeAg seronegative infections via IFN-I mediated 1896 preC G to A mutation and lowered HBV DNA or HBsAg production [76], suggesting the possibility of being more prone to mutation within the HBV genome in patients with rt269I of genotype C2 versus rt269L. IFN-I could mediate antiviral action via upregulating APOBEC3G that hypermutates the HBV genome [77]. In addition, in the previous study, we found that rt269I enhanced IFN-I-mediated APOBEC3G signaling in infected hepatocytes (Figure 5), which could lead to the higher mutation frequency in G to A mutation in preC and BCP in patients with rt269I [76]. Therefore, in this study, we address this issue and explore rt269L or rt269I signature mutations by analyzing mutation frequency and types from complete HBV genome sequences randomly selected from an HBV database—HBVdb (http://hbvdb.ibcp.fr (accessed on 14 March 2021)).

In this study, we found that overall mutation rates of HBV nucleotides and amino acids are significantly higher in rt269I HBV strains than in rt269L strains (Table 1 and Figure 2). Of note, dN and d of amino acid in Pol region or RT region were significantly higher in rt269I than in rt269L strains, suggesting that rt269I versus rt269L strains are more prone to HBV mutations, particularly NS mutations in HBV Pol region (Table 1 and Figure 3). Given the previous finding that HBV Pol mutations are significantly associated with lower HBV replication or HBeAg negative serostatus [68], higher HBV mutations found in rt269I types may be due to enhanced host immune responses after conversion into HBeAg seronegative status, which could result in persistent infections in chronic patients with genotype C2 infections.

Our data also showed that NS mutations within MHR or "a" determinant in the S region, a major target against host humoral immune response, are significantly higher in rt269I HBV strains than in rt269L strains (Table 2 and Figure 4), suggesting that rt269I versus rt269L strains are more likely to generate immune or vaccine escape variants, also contributing persistent infections of rt269I strains in chronic patients of genotype C2 infections.
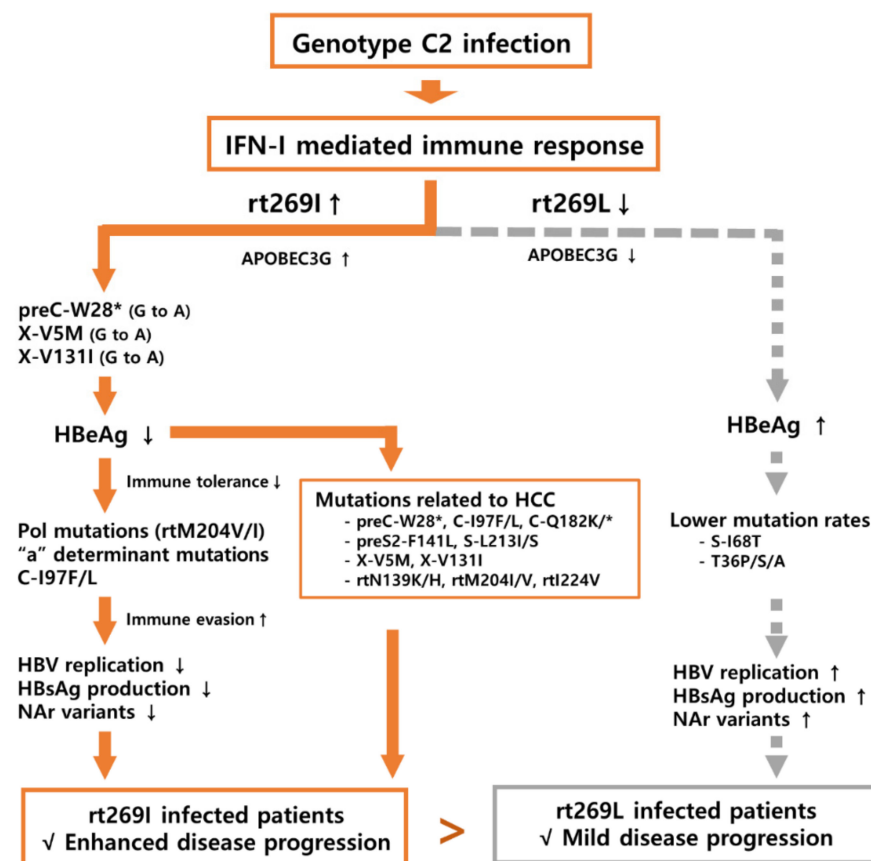
Figure 5. Schematic presentation. Schematic image showing that rt269I type versus rt269L type is more prone to overall genome mutations, particularly in the Pol and MHR regions in genotype C2 infections. The signature nonsynonymous mutations can contribute to the decreased level of HBV DNA replication, HBeAg secretion, NAr variants, and HCC occurrence. Finally, rt269I type versus rt269L type can lead to liver disease development.

In this study, we found a total of 19 NS signature mutation types, of which all 17 types, except for two types specific to rt269L (I68T in S region and T36P/S/A in X), were rt269I signature type. In rt269I signature NS mutations, several types, including preC-W28Stop (1896 preC mutation) and two types in the X region, V/L5M and V131I, are strongly related to HBeAg seronegative infections, which can lead to enhanced persistent rt269I type HBV infections by increasing mutation frequencies within the HBV genome, particularly in the Pol region, perhaps because of a host immune response. This could resultantly contribute to rt269I-type liver disease progression during the natural course of HBV genotype C2 infections (Figure 5). Indeed, we found several rt269I signature mutation types related to lower HBV replication (I84T/M/L/V in PreS1, preC-W28Stop, V/L5M in X region, C-I97F/L in C region and rtM204I/V in RT region) or liver disease progression (three types in preC/C region (preC-W28Stop, C-I97F/L and C-Q182K/*), F141L in PreS2, L213I/S in S region, three types in X region (V/L5M, T36P/S/A and V131I), three types in Pol region (rtN139K/H, rtM204I/V and rtI224V)) (Table 3 and Figure 5).

The present study has several major limitations. First, HBV genome sequences were used from only an HBV database, HBVdb (http://hbvdb.ibcp.fr (accessed on 14 March 2021)), and differences between distinct clades within sub-genotype C2 were not considered in this study. To elucidate further details regarding molecular epidemiologic traits of rt269L and rt269I strains in the chronic patients of genotype C2, a larger sample size of HBV sequences from multiple HBV databases reflecting diverse geo-ethnic traits and distinct clades within sub-genotype C2 has to be analyzed in future study. Second, although some meaningful differences between rt269L and rt269I strains were found from *t*-test-based analysis, the difference could not reach statistical significance in multiple comparisons

using Benjamini–Hochberg correction. Given that the variation between samples was rather large [39], it could be supplemented by further study using larger sample size of HBV sequences.

In conclusion, our data showed that rt269I type versus rt269L type is more prone to overall genome mutations, particularly in the Pol region in genotype C2 infections, and is more prevalent in signature NS mutations related to lowered HBV DNA replication, HBsAg and HBeAg secretion, potential NAr variants and HCC, perhaps via IFN-I mediated enhanced inflammation. Our data suggest that rt269I could contribute to liver disease progression via the generation of immune escape or potential NAr variants in chronic patients of genotype C2 (Figure 5).

Supplementary Materials: The following are available online at https://www.mdpi.com/2076-260 7/9/3/601/s1, Table S1. All the accession numbers of each sequence and the origin of the virus.

Author Contributions: H.J., D.H.K.: conceptualization, methodology, validation, investigation, visualization, and writing; Y.-M.C.: conceptualization methodology, validation, partial investigation, and visualization; H.C.: methodology, validation, statistical investigation; D.K.: validation, writing, supervision; B.-J.K.: conceptualization, project administration, supervision, writing—review and editing, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data produced here is available and can produced when required.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schweitzer, A.; Horn, J.; Mikolajczyk, R.T.; Krause, G.; Ott, 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. Yim, S.Y.; Kim, J.H. The epidemiology of hepatitis B virus infection in Korea. *Korean J. Intern. Med.* **2019**, *34*, 945–953. [CrossRef] [PubMed]
3. Liang, T.J.; Hepatitis, B. The virus and disease. *Hepatology* **2009**, *49*, S13–S21. [CrossRef] [PubMed]
4. Nowak, M.A.; Bonhoeffer, S.; Hill, A.M.; Boehme, R.; Thomas, H.C.; McDade, H. Viral dynamics in hepatitis B virus infection. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4398–4402. [CrossRef] [PubMed]
5. Kurbanov, F.; Tanaka, Y.; Mizokami, M. Geographical and genetic diversity of the human hepatitis B virus. *Hepatol. Res.* **2010**, *40*, 14–30. [CrossRef] [PubMed]
6. Chan, H.L.; Wong, M.L.; Hui, A.Y.; Hung, L.C.; Chan, F.K.; Sung, J.J. Hepatitis B virus genotype C takes a more aggressive disease course than hepatitis B virus genotype B in hepatitis B e antigen-positive patients. *J. Clin. Microbiol.* **2003**, *41*, 1277–1279. [CrossRef]
7. Osiowy, C.; Giles, E.; Trubnikov, M.; Choudhri, Y.; Andonov, A. Characterization of acute and chronic hepatitis B virus genotype in Canada. *PLoS ONE* **2015**, *10*, e0136074. [CrossRef]
8. Croagh, C.M.; Desmond, P.V.; Bell, S.J. Genotypes and viral variants in chronic hepatitis B: A review of epidemiology and clinical relevance. *World J. Hepatol.* **2015**, *7*, 289–303. [CrossRef]
9. Chan, H.L.; Hui, A.Y.; Wong, M.L.; Tse, A.M.; Hung, L.C.; Wong, V.W.; Sung, J.J. Genotype C hepatitis B virus infection is associated with an increased risk of hepatocellular carcinoma. *Gut* **2004**, *53*, 1494–1498. [CrossRef]
10. Kim, H.; Kim, B.J. Association of preS/S mutations with occult hepatitis B virus (HBV) infection in South Korea: Transmission potential of distinct occult HBV variants. *Int. J. Mol. Sci.* **2015**, *16*, 13595–13609. [CrossRef]
11. Kim, H.; Jee, Y.M.; Song, B.C.; Shin, J.W.; Yang, S.H.; Mun, H.S.; Kim, H.J.; Oh, E.J.; Yoon, J.H.; Kim, Y.J.; et al. Molecular epidemiology of hepatitis B virus (HBV) genotypes and serotypes in patients with chronic HBV infection in Korea. *Intervirology* **2007**, *50*, 52–57. [CrossRef]
12. Kim, H.; Jee, Y.M.; Song, B.C.; Hyun, J.W.; Mun, H.S.; Kim, H.J.; Oh, E.J.; Yoon, J.H.; Kim, Y.J.; Lee, H.S.; et al. Analysis of hepatitis B virus quasispecies distribution in a Korean chronic patient based on the full genome sequences. *J. Med. Virol.* **2007**, *79*, 212–219. [CrossRef]

13. Kim, B.J. Hepatitis B virus mutations related to liver disease progression of Korean patients. *World J. Gastroenterol.* **2014**, *20*, 460–467. [CrossRef]

14. Kim, D.W.; Lee, S.A.; Hwang, E.S.; Kook, Y.H.; Kim, B.J. Naturally occurring precore/core region mutations of hepatitis B virus genotype C related to hepatocellular carcinoma. *PLoS ONE* **2012**, *7*, e47372. [CrossRef] [PubMed]

15. Kim, H.; Lee, S.A.; Kim, D.W.; Lee, S.H.; Kim, B.J. Naturally occurring mutations in large surface genes related to occult infection of hepatitis B virus genotype C. *PLoS ONE* **2013**, *8*, e54486. [CrossRef]

16. Lee, S.A.; Kim, K.J.; Kim, D.W.; Kim, B.J. Male-specific W4P/R mutation in the pre-S1 region of hepatitis B virus, increasing the risk of progression of liver diseases in chronic patients. *J. Clin. Microbiol.* **2013**, *51*, 3928–3936. [CrossRef] [PubMed]

17. Kim, H.; Gong, J.R.; Lee, S.A.; Kim, B.J. Discovery of a novel mutation (X8Del) resulting in an 8-bp deletion in the hepatitis B virus X gene associated with occult infection in Korean vaccinated individuals. *PLoS ONE* **2015**, *10*, e0139551. [CrossRef] [PubMed]

18. Lee, S.A.; Kim, H.; Won, Y.S.; Seok, S.H.; Na, Y.; Shin, H.B. Male-specific hepatitis B virus large surface protein variant W4P potentiates tumorigenicity and induces gender disparity. *Mol. Cancer* **2015**, *14*, 23. [CrossRef]

19. Lee, S.A.; Kim, K.J.; Kim, H.; Choi, W.H.; Won, Y.S.; Kim, B.J. Hepatitis B virus preS1 deletion is related to viral replication increase and disease progression. *World J. Gastroenterol.* **2015**, *21*, 5039–5048. [CrossRef]

20. Song, B.C.; Kim, S.H.; Kim, H.; Ying, Y.H.; Kim, H.J.; Kim, Y.J.; Yoon, J.H.; Lee, H.S.; Cha, C.Y.; Kook, Y.H.; et al. Prevalence of naturally occurring surface antigen variants of hepatitis B virus in Korean patients infected chronically. *J. Med. Virol.* **2005**, *76*, 194–202. [CrossRef] [PubMed]

21. Song, B.C.; Kim, H.; Kim, S.H.; Cha, C.Y.; Kook, Y.H.; Kim, B.J. Comparison of full length sequences of hepatitis B virus isolates in hepatocellular carcinoma patients and asymptomatic carriers of Korea. *J. Med. Virol.* **2005**, *75*, 13–19. [CrossRef]

22. Kim, H.; Jee, Y.; Mun, H.S.; Park, J.H.; Yoon, J.H.; Kim, Y.J.; Lee, H.S.; Hyun, J.W.; Hwang, E.S.; Cha, C.Y.; et al. Characterization of two hepatitis B virus populations in a single Korean hepatocellular carcinoma patient with an HBeAg-negative serostatus: A novel X-Gene-deleted strain with inverted duplication sequences of upstream enhancer site II. *Intervirology* **2007**, *50*, 273–280. [CrossRef] [PubMed]

23. Kim, H.; Jee, Y.; Mun, H.S.; Song, B.C.; Park, J.H.; Hyun, J.W.; Hwang, E.S.; Cha, C.Y.; Kook, Y.H.; Kim, B.J. Comparison of full genome sequences between two hepatitis B virus strains with or without preC mutation (A1896) from a single Korean hepatocellular carcinoma patient. *Microbiol. Biotechnol.* **2007**, *17*, 701–704.

24. Kim, H.J.; Park, J.H.; Jee, Y.; Lee, S.A.; Kim, H.; Song, B.C.; Yang, S.; Lee, M.; Yoon, J.H.; Kim, Y.J.; et al. Hepatitis B virus X mutations occurring naturally associated with clinical severity of liver disease among Korean patients with chronic genotype C infection. *J. Med. Virol.* **2008**, *80*, 1337–1343. [CrossRef]

25. Mun, H.S.; Lee, S.A.; Jee, Y.; Kim, H.; Park, J.H.; Song, B.C.; Yoon, J.H.; Kim, Y.J.; Lee, H.S.; Hyun, J.W.; et al. The prevalence of hepatitis B virus preS deletions occurring naturally in Korean patients infected chronically with genotype C. *J. Med. Virol.* **2008**, *80*, 1189–1194. [CrossRef] [PubMed]

26. Lee, S.A.; Cho, Y.K.; Lee, K.H.; Hwang, E.S.; Kook, Y.H.; Kim, B.J. Gender disparity in distribution of the major hydrophilic region variants of hepatitis B virus genotype C according to hepatitis B e antigen serostatus. *J. Med. Virol.* **2011**, *83*, 405–411. [CrossRef] [PubMed]

27. Lee, S.A.; Mun, H.S.; Kim, H.; Lee, H.K.; Kim, B.J.; Hwang, E.S.; Kook, Y.H.; Kim, B.J. Naturally occurring hepatitis B virus X deletions and insertions among Korean chronic patients. *J. Med. Virol.* **2011**, *83*, 65–70. [CrossRef] [PubMed]

28. Mun, H.S.; Lee, S.A.; Kim, H.; Hwang, E.S.; Kook, Y.H.; Kim, B.J. Novel F141L pre-S2 mutation in hepatitis B virus increases the risk of hepatocellular carcinoma in patients with chronic genotype C infections. *J. Virol.* **2011**, *85*, 123–132. [CrossRef] [PubMed]

29. Kim, D.W.; Lee, S.A.; Kim, H.; Won, Y.S.; Kim, B.J. Naturally occurring mutations in the nonstructural region 5B of hepatitis C virus (HCV) from treatment-naïve Korean patients chronically infected with HCV genotype 1b. *PLoS ONE* **2014**, *9*, e87773. [CrossRef] [PubMed]

30. Hayer, J.; Jadeau, F.; Deléage, G.; Kay, A.; Zoulim, F.; Combet, C. HBVdb: A knowledge database for Hepatitis B Virus. *Nucleic Acids Res.* **2013**, *41*, D566–D570. [CrossRef] [PubMed]

31. Yin, Y.; He, K.; Wu, B.; Xu, M.; Du, L.; Liu, W.; Liao, P.; Liu, Y.; He, M. A systematic genotype and subgenotype re-ranking of hepatitis B virus under a novel classification standard. *Heliyon* **2019**, *5*, e02556. [CrossRef] [PubMed]

32. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef] [PubMed]

33. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [CrossRef]

34. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526.

35. Minin, V.N.; Dorman, K.S.; Fang, F.; Suchard, M.A. Dual multiple change-point model leads to more accurate recombination detection. *Bioinformatics* **2005**, *21*, 3034–3042. [CrossRef] [PubMed]

36. Fang, Z.L.; Sabin, C.A.; Dong, B.Q.; Wei, S.C.; Chen, Q.Y.; Fang, K.X.; Yang, J.Y.; Wang, X.Y.; Harrison, T.J. The association of HBV core promoter double mutations (A1762T and G1764A) with viral load differs between HBeAg positive and anti-HBe positive individuals: A longitudinal analysis. *J. Hepatol.* **2009**, *50*, 273–280. [CrossRef]

37. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]

38. Kwong, K.S.; Holland, B.; Cheung, S.H. A modified Benjamini-Hochberg multiple comparisons procedure for controlling the false discovery rate. *J. Stat. Plan. Inference* **2002**, *104*, 351–362. [CrossRef]
39. Ferreira, J.A.; Zwinderman, A.H. Approximate power and sample size calculations with the Benjamini-Hochberg method. *Int. J. Biostat.* **2006**, *2*, 8. [CrossRef]
40. Khodadad, N.; Seyedian, S.S.; Moattari, A.; Haghighi, S.B.; Pirmoradi, R.; Abbasi, S.; Makvandi, M. In silico functional and structural characterization of hepatitis B virus PreS/S-gene in Iranian patients infected with chronic hepatitis B virus genotype D. *Heliyon* **2020**, *6*, e04332. [CrossRef]
41. Littlejohn, M.; Davies, J.; Yuen, L.; Edwards, R.; Sozzi, T.; Jackson, K.; Cowie, B.; Tong, S.; Davis, J.; Locarnini, S. Molecular virology of hepatitis B virus, sub-genotype C4 in northern Australian Indigenous populations. *J. Med. Virol.* **2014**, *86*, 695–706. [CrossRef] [PubMed]
42. Shi, Y.; Wei, F.; Hu, D.; Li, Q.; Smith, D.; Li, N.; Chen, D. Mutations in the major hydrophilic region (MHR) of hepatitis B virus genotype C in North China. *J. Med. Virol.* **2012**, *84*, 1901–1906. [CrossRef]
43. Choi, Y.M.; Lee, S.Y.; Kim, B.J. Naturally Occurring Hepatitis B Virus Mutations Leading to Endoplasmic Reticulum Stress and Their Contribution to the Progression of Hepatocellular Carcinoma. *Int. J. Mol. Sci.* **2019**, *20*, 597. [CrossRef] [PubMed]
44. Datta, S.; Ghosh, A.; Dasgupta, D.; Ghosh, A.; Roychoudhury, S.; Roy, G.; Das, S.; Das, K.; Gupta, S.; Basu, K.; et al. Novel point and combo-mutations in the genome of hepatitis B virus-genotype D: Characterization and impact on liver disease progression to hepatocellular carcinoma. *PLoS ONE* **2014**, *9*, e110012. [CrossRef]
45. Chen, C.H.; Changchien, C.S.; Lee, C.M.; Hung, C.H.; Hu, T.H.; Wang, J.H.; Wang, J.C.; Lu, S.N. Combined mutations in pre-s/surface and core promoter/precore regions of hepatitis B virus increase the risk of hepatocellular carcinoma: A case-control study. *J. Infect. Dis.* **2008**, *198*, 1634–1642. [CrossRef] [PubMed]
46. Benhenda, S.; Cougot, D.; Buendia, M.A.; Neuveut, C. Chapter 4 hepatitis B virus X protein: Molecular functions and its role in virus life cycle and pathogenesis. *Adv. Cancer Res.* **2009**, *103*, 75–109. [CrossRef]
47. Quarleri, J. Core promoter: A critical region where the hepatitis B virus makes decisions. *World J. Gastroenterol.* **2014**, *20*, 425–435. [CrossRef]
48. Kim, H.; Lee, S.A.; Kim, B.J. X region mutations of hepatitis B virus related to clinical severity. *World J. Gastroenterol.* **2016**, *22*, 5467–5478. [CrossRef]
49. Schlicht, H.J.; Wasenauer, G. The quaternary structure, antigenicity, and aggregational behavior of the secretory core protein of human hepatitis B virus are determined by its signal sequence. *J. Virol.* **1991**, *65*, 6817–6825. [CrossRef]
50. Matsumura, S.; Yamamoto, K.; Shimada, N.; Okano, N.; Okamoto, R.; Suzuki, T.; Hakoda, T.; Mizuno, M.; Higashi, T.; Tsuji, T. High frequency of circulating HBcAg-specific CD8 T cells in hepatitis B infection: A flow cytometric analysis. *Clin. Exp. Immunol.* **2001**, *124*, 435–444. [CrossRef]
51. Jung, M.C.; Diepolder, H.M.; Spengler, U.; Wierenga, E.A.; Zachoval, R.; Hoffmann, R.M.; Eichenlaub, D.; Frösner, G.; Will, H.; Pape, G.R. Activation of a heterogeneous hepatitis B (HB) core and e antigenspecific CD4+ T-cell population during seroconversion to antiHBe and anti-HBs in hepatitis B virus infection. *J. Virol.* **1995**, *69*, 3358–3368. [CrossRef]
52. Yim, S.Y.; Um, S.H.; Young Jung, J.; Kim, T.H.; Kim, J.D.; Keum, B.; Seo, Y.S.; Yim, H.J.; Jeen, Y.T.; Lee, H.S.; et al. Clinical significance of hepatitis B virus precore and core promoter variants in Korean patients with chronic hepatitis B. *J. Clin. Gastroenterol.* **2015**, *49*, 61–68. [CrossRef]
53. Kim, H.; Lee, S.A.; Do, S.Y.; Kim, B.J. Precore/core region mutation of hepatitis B virus related to clinical severity. *World J. Gastroenterol.* **2016**, *22*, 4287–4296. [CrossRef] [PubMed]
54. Xie, Y.; Liu, S.; Zhao, Y.; Zhang, L.; Zhao, Y.; Liu, B. Precore/core region mutations in hepatitis B virus DNA predict postoperative survival in hepatocellular carcinoma. *PLoS ONE* **2015**, *10*, e0133393. [CrossRef]
55. Hosono, S.; Tai, P.C.; Wang, W.; Ambroso, M.; Hwang, D.G.; Yuan, T.T. Core antigen mutations of human hepatitis B virus in hepatomas accumulate in MHC class II-restricted T cell epitopes. *Virology* **1995**, *212*, 151–162. [CrossRef] [PubMed]
56. Yuan, T.T.T.; Sahu, G.K.; Whitehead, W.E.; Greenberg, R.; Shih, C. The mechanism of an immature secretion phenotype of a highly frequent naturally occurring missense mutation at codon 97 of human hepatitis B virus core antigen. *J. Virol.* **1999**, *73*, 5731–5740. [CrossRef]
57. Shirvani-Dastgerdi, E.; Winer, B.Y.; Celia-Terrassa, T.; Kang, Y.; Tabernero, D.; Yagmur, E.; Rodríguez-Frías, 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]
58. Thedja, M.D.; Muljono, D.H.; Ie, S.I.; Sidarta, E.; Turyadi; Verhoef, J.; Marzuki, S. Genogeography and immune epitope characteristics of hepatitis B virus genotype C reveals two distinct types: Asian and Papua-pacific. *PLoS ONE* **2015**, *10*, e0132533. [CrossRef]
59. Clark, D.N.; Flanagan, J.M.; Hu, J. Mapping of functional subdomain in the terminal protein domain of hepatitis B virus polymerase. *J. Virol.* **2016**, *91*, e01785-16. [CrossRef]
60. Li, X.G.; Liu, B.M.; Xu, J.; Liu, X.E.; Ding, H.; Li, T. Discrepancy of potential antiviral resistance mutation profiles within the HBV reverse transcriptase between nucleos(t)ide analogue-untreated and –treated patients with chronic hepatitis B in a hospital in China. *J. Med. Virol.* **2011**, *84*, 207–216. [CrossRef] [PubMed]

61. Lok, A.S.; Zoulim, F.; Locarnini, S.; Bartholomeusz, A.; Ghany, M.G.; Pawlotsky, J.M.; Liaw, Y.F.; Mizokami, M.; Kuiken, C. Hepatitis B virus drug resistance working group. Antiviral drug-resistant HBV: Standardization of nomeclature and assays and recommendations for management. *Hepatology* **2007**, *46*, 254–265. [CrossRef]

62. Locarnini, S. Primary resistance, multidrug resistance, and cross-resistance pathways in HBV as a consequence of treatment failure. *Hepatol. Int.* **2008**, *2*, 147–151. [CrossRef] [PubMed]

63. Ko, C.; Shin, Y.C.; Park, W.J.; Kim, S.; Kim, J.; Ryu, W.S. Residues Arg703, Asp777, and Arg781 of the RNase H domain of hepatitis B virus polymerase are critical for viral DNA synthesis. *J. Virol.* **2014**, *88*, 154–163. [CrossRef] [PubMed]

64. Xie, Y.; Liu, S.; Zhao, Y.; Guo, Z.; Xu, J. X protein mutations in hepatitis B virus DNA predict postoperative survival in hepatocellular carcinoma. *Tumor. Biol.* **2014**, *35*, 10325–10331. [CrossRef] [PubMed]

65. Asim, M.; Malik, A.; Sarma, M.P.; Polipalli, S.K.; Begum, N.; Ahmad, I.; Khan, L.A.; Husain, S.A.; Akhtar, N.; Husain, S. Hepatitis B virus BCP, Precore/core, X gene mutations/genotypes and the risk of hepatocellular carcinoma in India. *J. Med. Virol.* **2010**, *82*, 1115–1125. [CrossRef] [PubMed]

66. Preikschat, P.; Gunther, S.; Reinhold, S.; Will, H.; Budde, K.; Neumayer, H.H.; Kruger, D.H.; Meisel, H. Complex HBV populations with mutations in core promoter, C gene, and pre-S region are associated with development of cirrhosis in long-term renal transplant recipients. *Hepatology* **2002**, *35*, 466–477. [CrossRef]

67. Zhao, Y.; Wu, J.; Sun, L.; Liu, G.; Li, B.; Zheng, Y.; Li, X.; Tao, J. Prevalence of mutations in HBV DNA polymerase gene associated with nucleos(t)ide resistance in treatment-naïve patients with chronic hepatitis B in central China. *Braz. J. Infect. Dis.* **2016**, *20*, 173–178. [CrossRef]

68. Kim, J.E.; Lee, S.Y.; Kim, H.; Kim, K.J.; Choe, W.H.; Kim, B.J. Naturally occurring mutations in the reverse transcriptase region of hepatitis B virus polymerase from treatment-naïve Korean patients infected with genotype C2. *World J. Gastroenterol.* **2017**, *23*, 4222–4232. [CrossRef]

69. Yuen, L.K.; Locarnini, S. Genetic variability of hepatitis B virus and response to antiviral treatments: Searching for a bigger picture. *J. Hepatol.* **2009**, *50*, 445–448. [CrossRef]

70. Zheng, J.; Zeng, Z.; Zhang, D.; Yu, Y.; Wang, F.; Pan, C.Q. Prevalence and significance of hepatitis B reverse transcriptase mutants in different disease stages of untreated patients. *Liver Int.* **2012**, *32*, 1535–1542. [CrossRef]

71. 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 treatmentnaïve Chinese patients. *Antivir. Res.* **2010**, *85*, 512–519. [CrossRef] [PubMed]

72. Wang, L.P.; Han, F.Z.; Duan, H.L.; Ji, F.; Yan, X.B.; Fan, Y.C.; Wang, K. Hepatitis B virus pre-existing drug resistant mutation is related to the genotype and disease progression. *J. Infect. Dev. Ctries* **2017**, *11*, 727–732. [CrossRef] [PubMed]

73. Fan, J.; Zhang, Y.; Xiong, H.; Wang, Y.; Guo, X. Nucleotide analogueresistant mutations in hepatitis B viral genomes found in hepatitis B patients. *J. Gen. Virol.* **2015**, *96*, 663–670. [CrossRef]

74. Lin, C.L.; Kao, J.H. The clinical implications of hepatitis B virus genotype: Recent advances. *J. Gastroenterol. Hepatol.* **2011**, *26*, 123–130. [CrossRef] [PubMed]

75. Tang, L.S.Y.; Covert, E.; Wilson, E.; Kottilil, S. Chronic hepatitis B infection: A review. *JAMA* **2018**, *319*, 1802–1813. [CrossRef] [PubMed]

76. Lee, S.Y.; Choi, Y.M.; Oh, S.J.; Yang, S.B.; Lee, J.H.; Choe, W.H.; Kook, Y.H.; Kim, B.J. rt269I type of hepatitis B virus (HBV) leads to HBV e antigen negative infections and liver disease progression via mitochondrial stress mediated type I interferon production in chronic patients with genotype C infections. *Front. Immunol.* **2019**, *10*, 1735. [CrossRef]

77. Noguchi, C.; Hiraga, N.; Mori, N.; Tsuge, M.; Imamura, M.; Takahashi, S.; Fujimoto, Y.; Ochi, H.; Abe, H.; Maekawa, T.; et al. Dual effect of APOBEC3G on hepatitis B virus. *J. Gen. Virol.* **2007**, *88 Pt 2*, 432–440. [CrossRef]