Research article

JHEP|Reports

Whole genome deep sequencing analysis of viral quasispecies diversity and evolution in HBeAg seroconverters

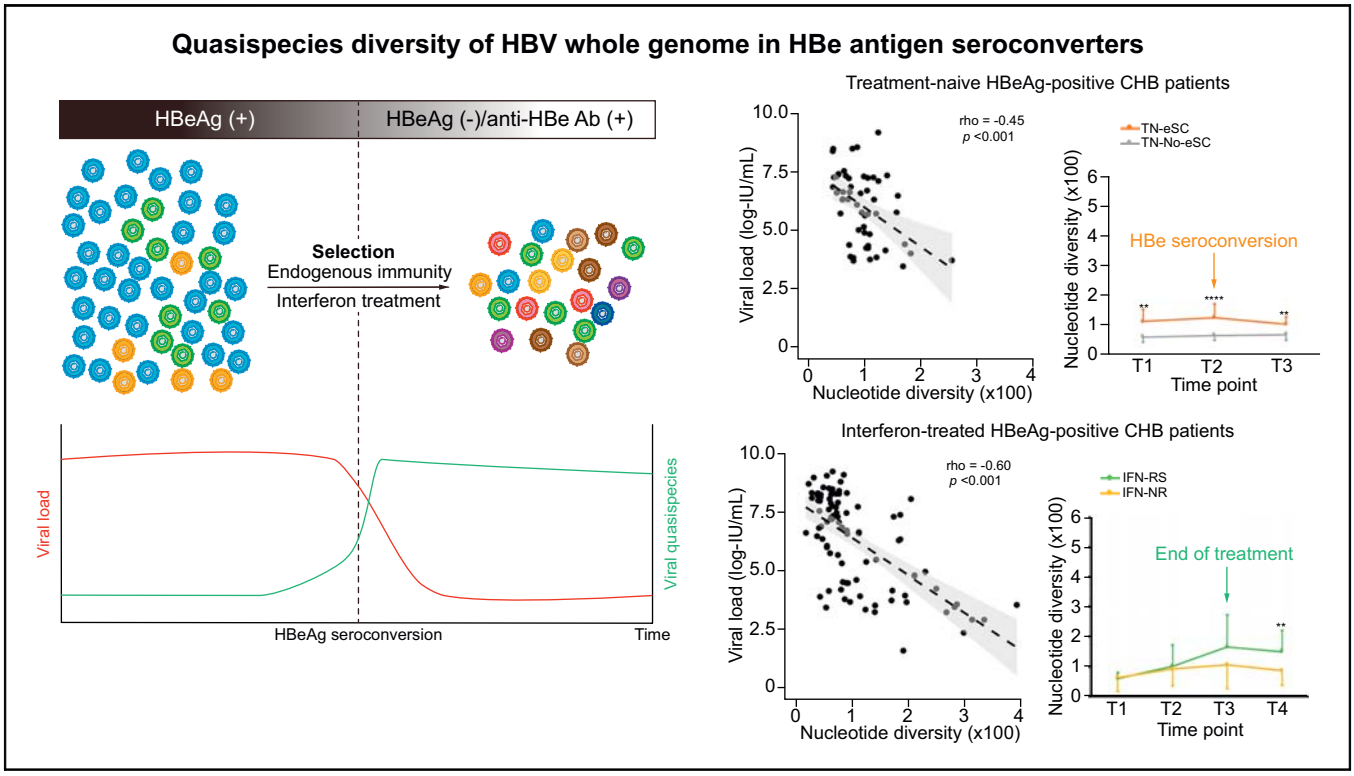
Authors

Su-Ru Lin, Ta-Yu Yang, Cheng-Yuan Peng, You-Yu Lin, Chia-Yen Dai, Hurng-Yi Wang, Tung-Hung Su, Tai-Chung Tseng, I-Jung Liu, Huei-Ru Cheng, Yueh-Chi Shen, Fang-Yi Wu, Chun-Jen Liu, Ding-Shinn Chen, Pei-Jer Chen, Hung-Chih Yang, Jia-Horng Kao

Correspondence

kaojh@ntu.edu.tw (J.-H. Kao), hcyang88@ntu.edu.tw (H.-C. Yang)

Graphical abstract



Highlights

- Deep sequencing of whole HBV genome uncovers the quasispecies changes in chronic hepatitis B patients.
- The nucleotide diversity of HBV negatively correlates with viraemia during HBeAg loss/seroconversion.
- Viral quasispecies diversity is greater in spontaneous HBeAg seroconverters before and after seroconversion than in treatment-naïve non-seroconverters.
- Responders to IFN have greater viral quasispecies diversity than non-responders at 24 weeks after treatment.
- The genome positions of non-synonymous intra-host single nucleotide variants (iSNVs) of HBV tend to be located at possible T cell epitopes.

Lay summary

HBeAg seroconversion is a landmark in the natural history of chronic HBV infection. Using next-generation sequencing, we found that the nucleotide diversity of HBV was negatively correlated with viral load and hepatitis activity. Patients undergoing HBeAg seroconversion had more diverse HBV genomes and a faster viral evolution rate. Our findings suggest HBeAg seroconversion is driven by host selection pressure, likely immune selection pressure.

https://doi.org/10.1016/j.jhepr.2021.100254

Research article

JHEP|Reports

Whole genome deep sequencing analysis of viral quasispecies diversity and evolution in HBeAg seroconverters

Su-Ru Lin,[1,†] Ta-Yu Yang,[1,†] Cheng-Yuan Peng,[2,3] You-Yu Lin,[4] Chia-Yen Dai,[5,6] Hurng-Yi Wang,[4] Tung-Hung Su,[7,8] Tai-Chung Tseng,[7,8] I-Jung Liu,[9] Huei-Ru Cheng,[4] Yueh-Chi Shen,[1] Fang-Yi Wu,[1] Chun-Jen Liu,[4,7,8,10,11] Ding-Shinn Chen,[4,7,8,10,12] Pei-Jer Chen,[4,7,8,10,11] Hung-Chih Yang,[1,4,7,8,*] Jia-Horng Kao[4,7,8,10,11,*]

[1]*Department of Microbiology, National Taiwan University, Taipei, Taiwan;* [2]*School of Medicine, China Medical University, Taichung, Taiwan;* [3]*Department of Internal Medicine, Center for Digestive Medicine, China Medical University Hospital, Taichung, Taiwan;* [4]*Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan;* [5]*Department of Internal Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan;* [6]*Hepato-Biliary Division, Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan;* [7]*Division of Gastroenterology and Hepatology, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan;* [8]*Hepatitis Research Center, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan;* [9]*Cardinal Tien Junior College of Healthcare and Management, New Taipei City, Taiwan;* [10]*Department of Internal Medicine, National Taiwan University, Taipei, Taiwan;* [11]*Department of Medical Research, National Taiwan University Hospital, Taipei, Taiwan;* [12]*Genomics Research Center, Academia Sinica, Taipei, Taiwan*

JHEP Reports **2021.** https://doi.org/10.1016/j.jhepr.2021.100254

Background & Aims: We aimed to investigate how viral quasispecies of the HBV whole genome evolves and diversifies in response to HBeAg seroconversion and viral control utilising next-generation sequencing (NGS).
Methods: Fifty HBeAg-positive chronic hepatitis B patients, including 18 treatment-naïve and 32 interferon (IFN)-treated individuals, were recruited. Serial HBV whole genomes in serum were analysed by NGS to determine sequence characteristics and viral quasispecies.
Results: HBV quasispecies diversity, measured by nucleotide diversity, was negatively correlated with viral load and hepatitis activity. Spontaneous HBeAg seroconverters exhibited significantly greater viral quasispecies diversity than treatment-naïve non-seroconverters from >1 year before seroconversion (0.0112 *vs.* 0.0060, p <0.01) to >1 year after seroconversion (0.0103 *vs.* 0.0068, p <0.01). IFN-induced HBeAg seroconverters tended to have higher viral genetic diversity than non-seroconverters along with treatment. Particularly, the IFN responders, defined as IFN-induced HBeAg seroconversion with low viraemia, exhibited significantly greater genetic diversity of whole HBV genome at 6 months post-IFN treatment than IFN non-responders (0.0148 *vs.* 0.0106, p = 0.048). Moreover, spontaneous HBeAg seroconverters and IFN responders exhibited significantly higher evolutionary rates and more intra-host single-nucleotide variants. Interestingly, in spontaneous HBeAg seroconverters and IFN responders, there were distinct evolutionary patterns in the HBV genome.
Conclusions: Higher HBV quasispecies diversity is associated with spontaneous HBeAg seroconversion and IFN-induced HBeAg seroconversion with low viraemia, conferring a favourable clinical outcome.
Lay summary: HBeAg seroconversion is a landmark in the natural history of chronic HBV infection. Using next-generation sequencing, we found that the nucleotide diversity of HBV was negatively correlated with viral load and hepatitis activity. Patients undergoing HBeAg seroconversion had more diverse HBV genomes and a faster viral evolution rate. Our findings suggest HBeAg seroconversion is driven by host selection pressure, likely immune selection pressure.
© 2021 The Author(s). Published by Elsevier B.V. on behalf of European Association for the Study of the Liver (EASL). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

HBeAg seroconversion is a landmark in the natural history of chronic HBV infection.[1] It also serves a surrogate serological endpoint of antiviral treatment for HBeAg-positive chronic hepatitis B (CHB) patients.[2] The underlying mechanisms that drive this transition are not completely understood. It is generally believed that HBeAg seroconversion results from complex interplays between virus and host immunity.[3]

HBV possesses an error-prone polymerase, and exhibits high replication and mutation rates, leading to a heterogeneous viral

Keywords: Chronic hepatitis B; HBeAg seroconversion; Intra-host single nucleotide variants.
Received 23 December 2020; received in revised form 26 January 2021; accepted 29 January 2021; available online 18 February 2021
† These authors contributed equally.
* Corresponding authors. Addresses: National Taiwan University, No. 7 Chung San South Road, Taipei 10002, Taiwan; Tel.: +886 2 312 3456 #67307; fax:+886 2 2382 5962 (J.-H. Kao); National Taiwan University No. 1, Sec. 1, Ren-Ai Road, Taipei 10051, Taiwan; Tel.: +886 2 312 3456 #88299; fax:+886 2 2391 5293 (H.-C. Yang).
E-mail addresses: kaojh@ntu.edu.tw (J.-H. Kao), hcyang88@ntu.edu.tw (H.-C. Yang).

population named quasispecies, within an infected individual.[4] The study of viral quasispecies dynamics and evolution is important to understand the adaptability and pathogenesis and to design better therapeutic strategies.[5–7] Therefore, investigating how HBV evolves along with HBeAg seroconversion should provide mechanistic insights into the interactions between virus and host.

Interferon (IFN) is currently approved for antiviral treatment of CHB.[8,9] It suppresses HBV replication through both direct antiviral and indirect immunomodulatory effects. HBeAg seroconversion occurs in approximately one-third of HBeAg-positive CHB patients treated by conventional and pegylated-IFNs.[10–12] Several naturally occurring HBV mutants, namely, precore and basal core promoter (BCP) mutations, are associated with HBeAg seroconversion.[13,14] A recent study has shown that the regions of the HBV genome with increasing viral diversity tend to occur within T- and B-cell epitopes during the immune clearance phase.[15] Consistently, previous studies analysed the evolution of the HBV precore/core region during HBeAg seroconversion using the clonal sequencing and demonstrated that patients undergoing HBeAg seroconversion exhibited significantly higher genetic diversity of the precore/core region.[16,17] The data also showed that the genetic diversity associated with the precore mutations in individuals who had HBeAg seroconversion. However, because only the precore/core region was studied in prior reports, dynamic changes of the entire HBV genome remain largely unclear.

The modern sequencing technology named next-generation sequencing (NGS) has enabled high-throughput analysis of thousands of sequences and is a powerful tool for characterising genetic diversity in a huge number of viral strains.[18] Therefore, NGS is particularly suitable to study a heterogeneous population of viruses with the nature of quasispecies, like HBV, within an infected individual.[19–21] Understanding the viral evolution and quasispecies diversity of the entire HBV genome in depth will help decipher the underlying regulatory mechanisms of HBeAg loss/seroconversion.

Materials and methods

Study participants and sampling

A total of 50 HBeAg-positive CHB patients who were infected by genotype B of HBV were enrolled in this study, including 18 treatment-naïve and 32 IFN-treated individuals. Among 18 treatment-naïve patients, 12 had undergone spontaneous HBeAg seroconversion and the remaining 6 age-matched non-seroconverters were recruited from a historical cohort.[14,22] Also, 20 patients received IFN-α-2b and the other 12 received pegylated-IFN-α-2b for 24 or 48 weeks. Among the 32 patients with IFN-based therapy, 17 patients had undergone HBeAg seroconversion. This study was approved by the Research Ethical Committee of the National Taiwan University Hospital, China Medical University Hospital, and Kaohsiung Medical University Hospital.

Sera from the treatment-naïve HBeAg seroconverters were collected approximately 1–2 years before HBeAg seroconversion, at HBeAg seroconversion, and 0.5–3 years after HBeAg seroconversion. Sera from the age-matched treatment-naïve non-seroconverters were also collected at 3 comparable time points. For the IFN-treated patients, sera were collected at baseline (before treatment), at 3 months on treatment, at the end of treatment (EOT), and around 6 months after EOT.

DNA extraction and library construction

Viral DNA was isolated from 200 μl serum using a QIAamp DNA Blood Mini Kit (Qiagen, Germany) according to the manufacturer's manual and was amplified by HBV specific primer pairs (Supplementary CTAT Table). The detailed strategy for library construction is described in the Supplementary methods and Figure S1A. Only patients with at least 2 successful analyses of samples, including baseline samples, were included in this study.

Data processing and analysis

The raw compressed FASTQ files were generated by the build-in software in the MiSeq Sequencer (Illumina). The detailed procedures for data processing and analysis are described in the Supplementary methods.

Results

Demographic profiles of patients

A total of 50 HBeAg-positive CHB patients, including 18 treatment-naïve and 32 IFN-treated individuals, were recruited. We only enrolled patients who were infected by genotype B of HBV to minimise the effect of genotype difference. The clinical features of 50 study participants are shown in Table 1. Twelve of 18 treatment-naïve HBeAg-positive patients had undergone spontaneous HBeAg seroconversion, termed treatment-naïve HBeAg seroconverters (TN-eSC), and the other 6 age-matched treatment-naïve non-seroconverters (TN-No-eSC) who remained positive HBeAg during the follow-up period served as controls. The viral loads at baseline (1–3 years before HBeAg seroconversion) and at 1–3 years after HBeAg seroconversion in the TN-No-eSC group were significantly higher than those in the TN-eSC group. Among the 32 IFN-treated HBeAg-positive CHB patients, 17 were HBeAg seroconverters (IFN-eSC) and the remaining 15 patients were non-seroconverters (IFN-No-eSC) at 6 months after EOT. The baseline (pretreatment) alanine aminotransferase (ALT) level in the IFN-eSC group was significantly higher than that in the IFN-No-eSC group. The IFN-No-eSC group had significantly higher viral load at 6 months after EOT than the IFN-eSC group.

Deep sequencing of the HBV whole genomes

We chose the Illumina MiSeq platform using the 2×300 bp paired-end sequencing to analyse HBV genomes because it yields

Table 1. Clinical features of study cohort with treatment naïve or IFN treatment.

	Treatment-naïve		IFN-treated	
	eSC	No-eSC	eSC	No-eSC
Patients, n	12	6	17	15
Sex (M/F)	7/5	5/1	13/4	12/3
Age	29.8 ± 7.0	29.2 ± 7.8	29.0 ± 6.0	30.3 ± 7.7
ALT (U/L)				
Baseline	52.1 ± 40.1	137.0 ± 193.8	185.1 ± 81.3*	115.7 ± 72.7
Endpoint	46.9 ± 39.0	61.3 ± 23.0	44.2 ± 56.2	81.7 ± 95.0
Log-HBV DNA (IU/ml)				
Baseline	5.7 ± 1.6	7.9 ± 0.6*	7.3 ± 1.0	7.7 ± 0.7
Endpoint	5.4 ± 1.7	7.0 ± 1.0*	4.0 ± 2.0	6.1 ± 2.2*

ALT, alanine aminotransferase; EOT, end of treatment; eSC, HBeAg seroconverters; No-eSC, HBeAg non-seroconverters; IFN, interferon.
* Significant difference (p <0.05 by Mann-Whitney U test) was noted between the eSC and No-eSC groups. Endpoint: for treatment-naïve patients, the endpoint is 2–3 years after HBeAg seroconversion; for IFN-treated patients, the endpoint is at 6 months after EOT.

millions of relatively longer sequencing reads with lower error rate in a high-throughput single sequencing run. To make NGS sequencing cost-effective, we utilised a PCR-based library preparation method to generate barcode-tagged amplicons, which were subsequently pooled for MiSeq sequencing in a multiplexed indexing manner. The detailed sequencing and data processing strategies are described in Fig. S1. The median depth of coverage was 18,793×, ranging from 5,342× to 40,677×. The plasmid with HBV genotype B insert was also analysed by this NGS approach to determine the strategical error rate as average of minor allele percentage per site. The estimated error rate of this NGS approach was about 0.1%.

To validate our analysing strategy, we compared the frequencies of the precore and BCP mutations measured by NGS in this study to those measured by pyrosequencing in our previous study.[14] The results showed these 2 measurements were well correlated, and the Pearson's correlation coefficients (r) were 0.90, 0.90, and 0.84 for the measurements of A1762T, G1764A, and G1896A, respectively (Fig. S2).

Inverse correlation of HBV genetic diversity levels with viral load and ALT levels

It has been suggested that viral quasispecies reflects the viral adaptation to the host selection pressure and is associated with the clinical outcomes of viral infections. We thus determined the quasispecies diversity of HBV by calculating the average nucleotide diversity per site, which represents the heterogeneity of viral quasispecies.[23] We also compared the average nucleotide diversity to the average Shannon entropy per site and demonstrated these 2 measurements were highly correlated (Pearson's r = 0.98, Fig. S3).

To elucidate the relationship between HBV quasispecies diversity and host control of viruses, we compared the nucleotide diversity level with serum viral load (Fig. 1 and S4), which represents levels of viral replication. Interestingly, the nucleotide diversity levels were inversely correlated with viral loads among treatment-naïve (Spearman's correlation coefficient rho = -0.45, p <0.001, Fig. 1A) and IFN-treated (rho = -0.60, p <0.001, Fig. 1C) patients. This correlation was even stronger in IFN-eSC than that in IFN-No-eSC (rho = -0.70 *vs.* rho = -0.50) (Fig. S4E and F). We also determined the relationship between nucleotide diversity and ALT, which is considered an indirect indicator of immune-mediated liver injury in CHB patients. There was also significant inverse correlation between nucleotide diversity levels and ALT levels among treatment-naïve (rho = -0.30, p = 0.029, Fig. 1B) and IFN-treated (rho = -0.53, p <0.001, Fig. 1D) patients, respectively. Taken together, it implies that quasispecies diversification of HBV correlates with low viraemia and high hepatitis activity, suggesting viral quasispecies diversity may reflect the virological control by host immune-mediated cytolytic activity.

Higher viral quasispecies diversity in spontaneous HBeAg seroconverters

Next, we investigated the dynamic change of HBV quasispecies diversity along with the course of HBeAg seroconversion at 3 time points, 1–3 years before (T1), at (T2), and 1–3 years after (T3) HBeAg seroconversion. The nucleotide diversity of the HBV whole genome in the TN-eSC group was significantly higher than that in the TN-No-eSC group at all 3 comparable time points, through >1 year before seroconversion (0.0112 *vs.* 0.0059, p <0.01) to >1 year after seroconversion (0.0103 *vs.* 0.0068, p <0.01) (Fig. 2A). Further analysis of the nucleotide diversity of 4 individual open reading frames (ORFs), including core (C), polymerase (P), surface (S), and X, also showed similar patterns of comparison between the TN-eSC and the TN-No-eSC groups (Fig. 2B). We further performed the receiver operating characteristics analysis to find out the optimal nucleotide diversity cut-off for predicting spontaneous HBeAg seroconversion. The area under curve (AUC) is 0.889 for the nucleotide diversity at T1. We found 2 optimal cut-offs of nucleotide diversity, 0.0073 and 0.0099, with 83% and 67% sensitivity and 83% and 100% specificity, respectively, for predicting the 2-year chance of spontaneous HBeAg seroconversion.

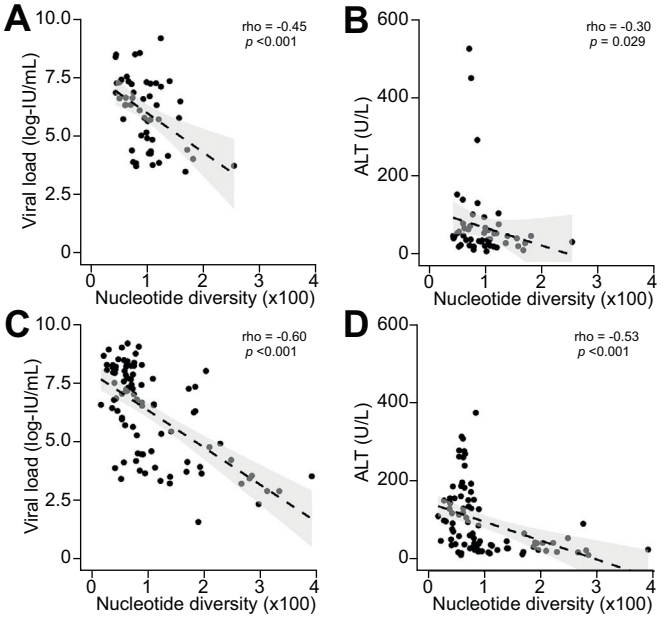


Fig. 1. The relationships between nucleotide diversity and the levels of viral load and ALT among the treatment-naïve and IFN-treated CHB patients. (A, B) The relationships between nucleotide diversity and the levels of viral load (A) and ALT (B) among the treatment-naïve CHB patients. (C, D) The relationships between nucleotide diversity and the levels of viral load (C) and ALT (D) among the IFN-treated patients. The relationship and significance are determined by Spearman correlation. ALT, alanine aminotransferase; CHB, chronic hepatitis B; IFN, interferon.

HBV has a compact genome with 4 extensively overlapped ORFs. It has been shown that the genetic diversity of the non-overlapping regions is higher than that of overlapping regions probably because the latter has a greater intrinsic restriction for viral mutations. We thus analysed the nucleotide diversity of the overlapping and non-overlapping regions in the HBV whole genome and individual viral genes except for the S gene, which is completely overlapped with the P gene. We found that in the core gene, the non-overlapping region had higher genetic diversity than the overlapping region in both TN-eSC and TN-No-eSC groups at T1 (Fig. S5).

We also measured the viral genetic diversity at amino acid levels. For fair comparison of the average amino acid heterogeneity per site in HBV genome, we analysed amino acids deduced from the non-overlapping regions within HBV whole genome and 4 ORFs, respectively. We found that the levels of amino acid diversity of the non-overlapping regions were significantly higher in the TN-eSC than in the TN-No-eSC group along with HBeAg seroconversion (Fig. 2C). Also, the levels of the amino acid

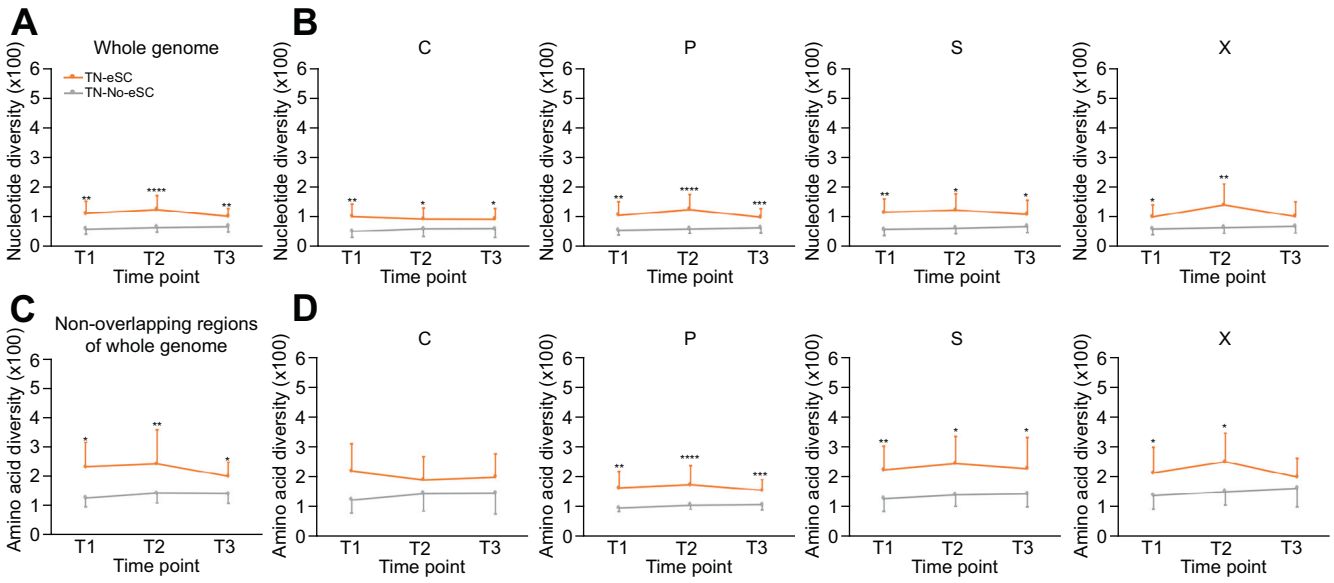


Fig. 2. The dynamic changes of HBV nucleotide and amino acid diversities along with spontaneous HBeAg seroconversion. (A) and (B) represent the nucleotide diversities of HBV whole genome, C, P, S, and X genes, respectively, before (T1), at (T2), and after (T3) the year of HBeAg seroconversion in TN-eSC (orange) and TN-No-eSC (grey) groups. (C) and (D) represent the amino acid diversities of the non-overlapping regions within HBV whole genome, C, P, S, and X genes, respectively, before (T1), at (T2), and after (T3) the year of HBeAg seroconversion in TN-eSC (orange) and TN-No-eSC (grey) groups. The value is shown as the mean ± SD. Asterisks indicate the statistical significance in comparison between groups by Mann-Whitney *U* test (**p* <0.05, ***p* <0.01, ****p* <0.005, *****p* <0.001). C, core; P, polymerase; S, surface; TN-eSC, treatment-naïve HBeAg seroconverters; TN-No-eSC, treatment-naïve non-seroconverters.

diversity of 4 ORFs were higher in the TN-eSC group than those in the TN-No-eSC group (Fig. 2D).

Higher viral quasispecies diversity in IFN-treatment responders

In the IFN-treated HBeAg-positive CHB patients, the quasispecies diversity of HBV was measured at 4 serial time points, before treatment (T1), at 3 months on IFN treatment (T2), at EOT (T3), and at 6 months after EOT (T4). We found that IFN-treated HBeAg seroconverters (IFN-eSC) also tended to exhibit higher nucleotide diversity of the whole HBV genome and of 4 ORFs than non-seroconverters (IFN-No-eSC), but the difference was not as significant as that between TN-eSC and TN-No-eSC groups (Fig. S6). It is known that some IFN-treated CHB patients develop HBeAg seroconversion as late as at 48 weeks off-treatment.[24] In addition, viral control is also an important index for the treatment response to IFN. Thus, we stratified IFN-treated patients into 19 IFN-responders (IFN-RS) who are defined as HBeAg seroconversion with low viraemia (serum HBV DNA <20,000 IU/ml) at 48 weeks off-treatment, and 13 IFN-non-responders (IFN-NR), who did not meet the above criteria. We found that compared with the IFN-NR group, the IFN-RS group at T4 had significantly higher nucleotide diversity of HBV whole genome (0.0151 *vs.* 0.0086, $p = 0.0052$) and of the core (0.0126 *vs.* 0.0086, $p = 0.0153$) and polymerase (0.0152 *vs.* 0.0085, $p = 0.0069$) genes (Fig. 3A and B). To determine whether nucleotide diversity could be used to predict IFN response, we further measured the AUC for the nucleotide diversity at T2 and T3. However, neither of them provided good predictive ability for IFN response (AUC was 0.507, 95% CI 0.286–0.728, for nucleotide diversity at T2; AUC was 0.682, 95% CI 0.476–0.888, for nucleotide diversity at T3).

In addition, we compared the nucleotide diversity levels between the non-overlapping and overlapping regions. The data showed that the non-overlapping region had significantly higher genetic diversity than the overlapping region only in the C gene of IFN-RS (Fig. S7). Collectively, these findings suggest that the C gene is the potential target of immunologic pressure induced by IFN treatment.

The viral genetic diversity at the amino acid level was also compared between the IFN-RS and IFN-NR groups. The amino acid heterogeneity in the non-overlapping regions of HBV whole genome was significantly higher in the IFN-RS group than that in the IFN-NR group at T4 (Fig. 3C). The patterns of amino acid heterogeneity was also similar to those of the nucleotide diversity of 4 ORFs although the difference between the 2 groups IFN-RS and IFN-NR was not significant (Fig. 3D).

Higher negative selection pressure in spontaneous HBeAg seroconverters and IFN responders

The ratio of dN/dS, where nonsynonymous substitution rate (dN) is divided by synonymous substitution rate (dS), is often used to measure the selection pressure for the evolution of individual genes. To probe the selection pressure on individual viral genes, we thus calculated the dN/dS ratio of each viral ORF. We found that spontaneous HBeAg seroconverters had significantly lower dN/dS of surface protein than non-seroconverters (Fig. 4A). In addition, the IFN responders had significantly lower dN/dS ratio of core protein than non-responders (Fig. 4B). This indicates that there was indeed the negative selective pressure exerted over viral proteins.

Higher evolutionary rate of HBV genome in spontaneous HBeAg seroconverters and IFN-treatment responders

In addition to analysing the viral quasispecies diversity, we also computed the evolutionary rate of the HBV whole genome, which is defined as the number of substitutions per site per month. In treatment-naïve HBeAg-positive patients, the

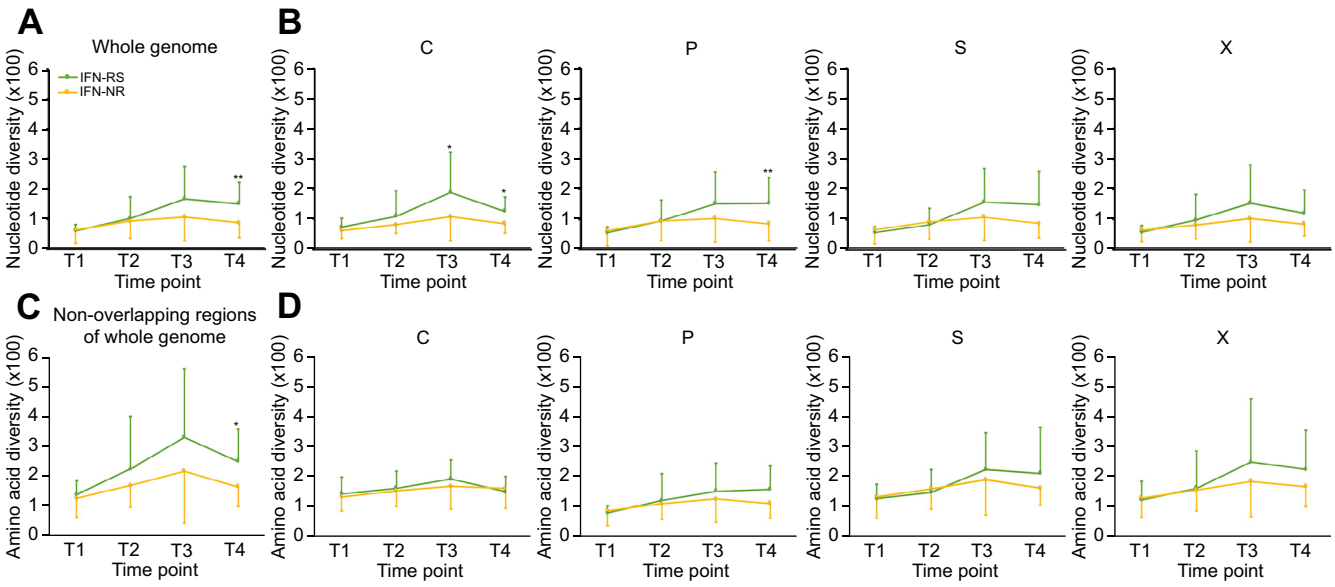


Fig. 3. The dynamic changes of HBV nucleotide and amino acid diversities along with IFN treatment. (A) and (B) represent the nucleotide diversities of HBV whole genome, C, P, S, and X ORFs, respectively, before (T1), at 12 weeks of (T2), at the end of (T3), and 24 weeks after (T4) IFN treatment in IFN-RS (green) and IFN-NR (yellow) groups. (C) and (D) represent the amino acid diversities of the non-overlapping regions within the HBV whole genome, C, P, S, and X ORFs, respectively, before (T1), at 12 weeks of (T2), at the end of (T3), and 24 weeks after (T4) IFN treatment in IFN-RS (green) and IFN-NR (yellow) groups. Values are shown as the mean ± SD. Asterisks indicate the statistical significance in comparison between groups by Mann-Whitney *U* test (**p* <0.05, ***p* <0.01). C, core; IFN, interferon; IFN-NR, IFN-non-responders; IFN-RS, IFN-responders; ORFs, open reading frames; P, polymerase; S, surface.

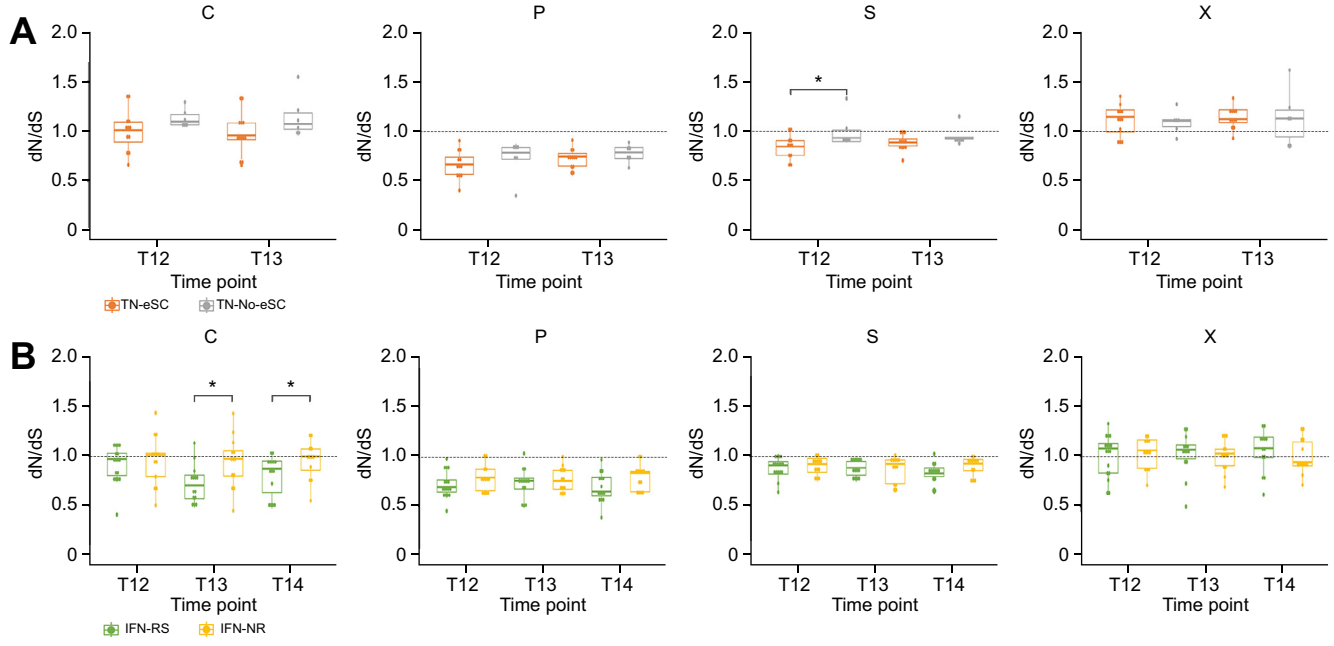


Fig. 4. The dN/dS ratios along with HBeAg seroconversion. (A) The dN/dS ratios of C, P, S, and X ORFs in the treatment-naïve groups. The colours, orange and grey, indicate the TN-eSC and the TN-No-eSC groups, respectively. (B) The dN/dS ratios of C, P, S, and X ORFs in the IFN-treated groups. The colours green and yellow indicate the IFN-RS and the IFN-NR groups, respectively. Values are shown as the median (IQR). Asterisks indicate the statistical significance by Mann-Whitney *U* test (****p* <0.005, *****p* <0.001). C, core; dN, nonsynonymous substitution rate; dS, synonymous substitution rate; IFN, interferon; IFN-NR, IFN-non-responders; IFN-RS, IFN-responders; ORFs, open reading frames; P, polymerase; S, surface; TN-eSC, treatment-naïve HBeAg seroconverters; TN-No-eSC, treatment-naïve non-seroconverters.

evolutionary rate of HBV whole genome in the TN-eSC group was significantly greater than that in the TN-No-eSC group (0.0011 *vs.* 0.0002 substitution/site/month before seroconversion, p <0.001; 0.0010 *vs.* 0.0002 substitution/site/month after seroconversion, p = 0.001) (Fig. 5A). In addition, the HBV evolutionary rate in the IFN-RS was also greater than that in the IFN-NR group after EOT (between T3 and T4) (0.0042 *vs.* 0.0020 substitution/site/month, p = 0.002) (Fig. 5B).

We sought to examine the genetic variation patterns of the genetic distances across the entire HBV genome between viral populations from baseline to the time after HBeAg seroconversion (Fig. S8). This allows us to scan the entire HBV genome under selection pressure. By the sliding window analysis of genetic distance, we compared the evolution patterns at three different scales, including between HBV-A, -B, -C, and -D genotypes, within genotype B, and within individuals among different defined groups. The genetic variation patterns between HBV genotypes and within genotype B were highly correlated (r = 0.75, p <0.001). Both the TN-eSC and IFN-RS groups had greater distance compared with the TN-No-eSC and IFN-NR groups. Interestingly, the genetic variation pattern of the IFN-RS group moderately but significantly correlated with that between HBV genotypes (r = 0.32, p <0.001) and with that within HBV-B (r = 0.29, p <0.001). In contrast, the genetic variation pattern of the IFN-NR did not positively correlate with those of between HBV-A, -B, -C, and -D (r = -0.02), and within HBV-B (r = -0.15). Furthermore, the genetic variation pattern of the TN-eSC group did not show positive correlation with those of between HBV-A, -B, -C, and -D and within HBV-B, either (r = -0.15 and -0.12, respectively). Surprisingly, although the TN-No-eSC group had significantly lower genetic distance than the TN-eSC group, the genetic variation pattern of the TN-No-eSC group correlated moderately with that of among HBV-A, -B, -C, and -D (r = 0.39, p <0.001), and weakly with that of within HBV-B (r = 0.14, p <0.001). It suggests that the evolutionary force to drive the diversification between and within HBV genotypes may also contribute to amplifying the variation in the treatment-naïve CHB (TN-No-eSC) and IFN-treatment responders (IFN-RS).

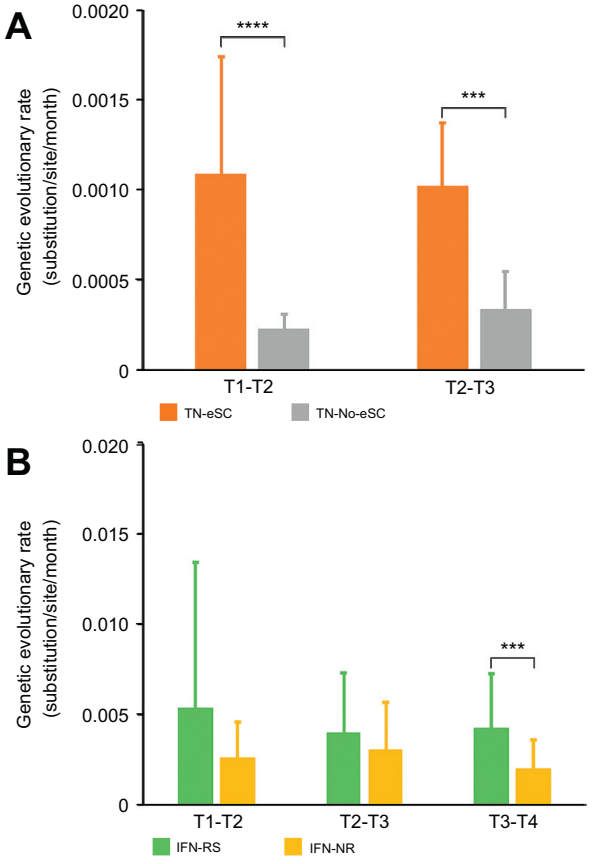


Fig. 5. The genetic evolutionary rate of HBV within the treatment-naïve and the IFN-treated HBeAg-positive patients. (A) The genetic evolutionary rate of HBV in the treatment-naïve groups. The colours orange and grey indicate the TN-eSC and the TN-No-eSC groups, respectively. (B) The genetic evolutionary rate of HBV in the IFN-treated groups. The colours green and yellow indicate the IFN-RS and the IFN-NR groups, respectively. Values are shown as the mean ± SD. Asterisks indicate the statistical significance by Mann-Whitney *U* test (***p <0.005, ****p <0.001). IFN, interferon; IFN-NR, IFN-non-responders; IFN-RS, IFN-responders; TN-eSC, treatment-naïve HBeAg seroconverters; TN-No-eSC, treatment-naïve non-seroconverters.

More HBV iSNVs in spontaneous HBeAg seroconverters and IFN-treatment responders

It has been evidenced that T cell immunity plays a major role in control viral replication in chronic HBV infection. We thus analysed the viral diversity within T cell epitopes. We hypothesised that viral genetic diversity should be higher within T cell epitopes than outside T cell epitopes. Because we did not have HLA profiles of patients in this study, we deduced the T cell epitopes based on the information of HLA allele frequency from the website Allele Frequency Net Database (http://www.allelefrequencies.net). We chose prevalent HLAs (>15% in population) in Taiwan, including HLA-A*02:01, HLA-A*02:03, HLA-A*02:07, HLA-A*11:01, HLA-A*24:02, HLA-B*40:01, HLA-B*46:01, HLA-C*01:02, and HLA-C*07:02, for predicting the cognate HBV epitopes by using the T cell epitope-prediction algorithms including SYFPEITHI,[25] Immune Epitope Database and analysis resource (IEDB),[26] BIMAS,[27] and EpiJen.[28] We chose 3–5 epitopes for each selected HLA allele with high scores across different algorithms, and in total got 9, 20, 12, and 8 epitopes within C, P, S, and X genes, respectively, which covers 1,602 nucleotides of the HBV genome. We then calculated the viral genetic diversity within and outside the predicted T cell epitopes. We found that TN-eSC and IFN-RS groups had higher viral genetic diversity within the predicted T cell epitopes than TN-No-eSC and IFN-NR, respectively. Similar patterns were also observed outside the predicted T cell epitopes. However, there was no significant difference of the genetic diversity between within and outside T cell epitopes that were identified through the above approach (Fig. S9).

The NGS approach also allows for examination of intra-host single nucleotide variants (iSNVs), which are defined as the minimum minor allele frequency of more than 5% within individual hosts.[29] Measuring the number of iSNVs also provides an opportunity to look into another aspect of viral quasispecies and adaptation. We observed that the numbers of iSNVs of the HBV whole genome and of the 4 ORFs in the TN-eSC group were significantly greater than those in the TN-No-eSC group (Fig. 6A and S10A). The average iSNVs per 100 bp of HBV whole genome in the TN-eSC were significantly greater than those in the TN-No-

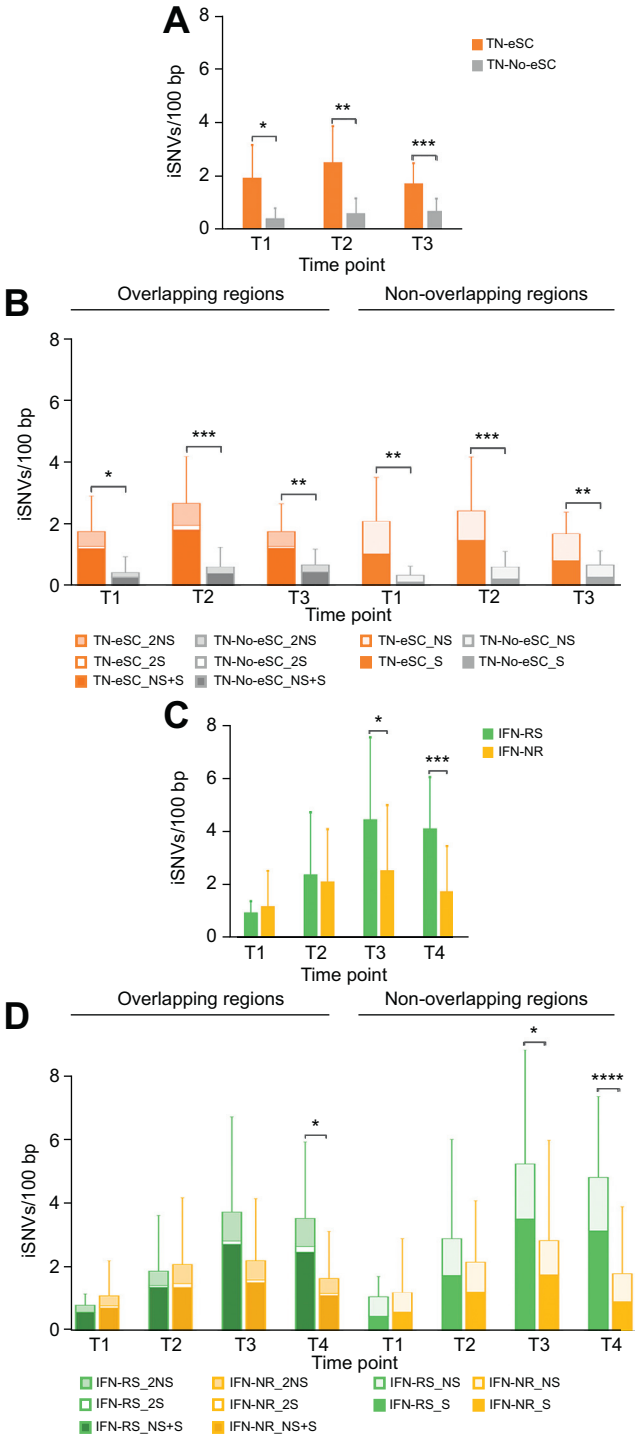


Fig. 6. The levels of iSNVs among the treatment-naïve and the IFN-treated HBeAg-positive CHB patients. (A) The normalised numbers of iSNVs in the HBV whole genome in the treatment-naïve groups. (B) The iSNVs attributes within the non-overlapping region of the HBV genome in the treatment-naïve groups. (C) The normalised numbers of iSNVs in the HBV whole genome in the IFN-treated groups. (D) The iSNVs attributes within the non-overlapping region of the HBV genome in the IFN-treated groups. NS represents the nonsynonymous substitution and S indicates the synonymous substitution. The number of iSNVs is normalised with the sequenced read length. Values are shown as the mean ± SD. The values of SD in panels (B) and (D) were computed by the SD derived from the normalised total iSNVs within the non-overlapping region. Asterisks indicate the statistical significance by Mann-Whitney *U* test (**p* <0.05, ***p* <0.01, ****p* <0.005, *****p* <0.001). CHB, chronic hepatitis B; IFN, interferon; IFN-NR, IFN-non-responders; IFN-RS, IFN-responders; iSNVs, intra-host single-nucleotide variants; TN-eSC, treatment-naïve HBeAg seroconverters; TN-No-eSC, treatment-naïve non-seroconverters.

eSC at all 3 time points (1.9 *vs.* 0.4 at T1, 2.5 *vs.* 0.6 at T2, and 1.7 *vs.* 0.7 at T3, all $p < 0.05$) (Fig. 6A). Likewise, the IFN-RS had significantly more iSNVs of the HBV whole genome than the IFN-NR at EOT (T3) (4.4 *vs.* 2.5 iSNVs/100 bp, $p = 0.0382$) and at 6 months after EOT (T4) (4.1 *vs.* 1.7 iSNVs/100 bp, $p = 0.0018$) (Fig. 6C). In addition, the IFN-RS also had significantly more iSNVs of C gene at both T3 (5.2 *vs.* 3.2 iSNVs/100 bp, $p = 0.034$) and T4 (3.7 *vs.* 1.7 iSNVs/100 bp, $p = 0.0005$), and more iSNVs of the P gene (4.1 *vs.* 1.7 iSNVs/100 bp, $p = 0.0017$), and the S gene (4.2 *vs.* 1.8 iSNVs/100 bp, $p = 0.0311$) at T4 (Fig. S10B).

Moreover, we determined the synonymous and non-synonymous substitutions of iSNVs in the overlapping and non-overlapping regions. However, measurement of the major non-synonymous changes of 2 overlapping genes is more complicated than that of a single non-overlapping gene. Therefore, we measured the synonymous and non-synonymous iSNVs for the overlapping and non-overlapping regions separately. We found that the TN-eSC group had significantly more iSNVs in the overlapping and non-overlapping regions than the TN-No-eSC group (Fig. 6B); the IFN-RS group also had significantly more iSNVs in the overlapping and non-overlapping regions than the IFN-NR group at both T3 and T4 (Fig. 6D).

We further identified iSNVs that had significantly differential frequency of the major allele between 2 distinct groups, which may represent the sites of HBV genome under the differential selection pressure leading to different clinical outcomes. The major allele at a particular site of HBV genome was defined by identifying the consensus (dominant) nucleotide allele at that position. To do so, we first calculated the nucleotide frequencies at each position based on the HBV sequences extracted from the HBV database (HBVdb, http://hbvdb.ibcp.fr/HBVdb/HBVdbIndex).[30] In total, there were 1,757 full-length HBV genotype B sequences, which were compiled from >300 filed isolated HBV genotype B strains. Interestingly, out of a total of 466, 490, and 425 iSNVs at T1, T2, and T3, respectively, we discovered 50, 31, and 23 iSNVs with significantly different major allele frequency between TN-eSC and TN-No-eSC groups (Table S1). Among them, we found 22, 16, and 16 iSNVs in the non-overlapping region, of which there were 2, 3, and 4 iSNVs at T1, T2, and T3, respectively, that had non-synonymous polymorphism between major and non-major (or minor) alleles, and were prevalent (>20% of individuals) in a particular group, either TN-eSC or TN-No-eSC (Table S1). In addition, we also found that 32, 76, 43, and 174 iSNVs with significantly different major allele frequency out of a total of 480, 700, 798, and 789 iSNVs between IFN-RS and IFN-NR groups at T1, T2, T3, and T4, respectively (Table S2). Among them, 1, 6, 1, and 2 iSNVs at T1, T2, T3, and T4, respectively, were prevalent and had non-synonymous polymorphism alleles in the non-overlapping regions (Table S3). Interestingly, except for the well-known naturally occurring precore mutants at nucleotide position 1,896, all the prevalent, non-synonymous iSNVs (nucleotide positions 1,762, 1,764, 2,304, 2,003, 1,126, 1,678, 1,726, 2,012, and 2,120) were located within the class I HLA epitopes (Table S3 and Fig. 7).[31–38] Taken together, the results strongly suggest that iSNVs emerged in response to the selection pressure of the host's T cell immunity.

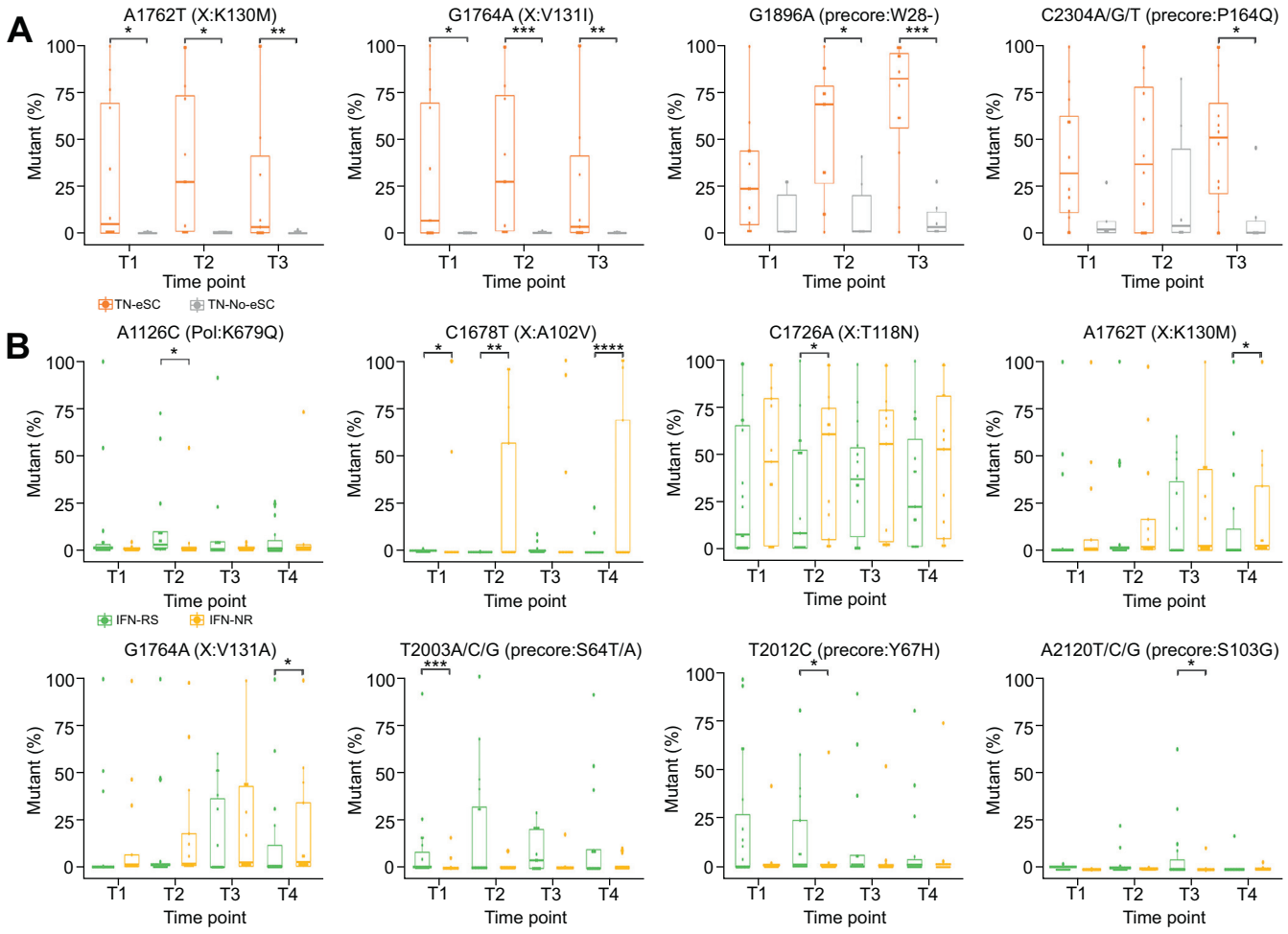


Fig. 7. The prevalently non-synonymous iSNVs with statistical importance in the non-overlapping regions of HBV genome. The prevalent iSNVs with non-synonymous substitutions and significantly different between (A) the treatment-naïve and (B) the IFN-treated groups. The nucleotide derived from the compiled HBV genotype B sequences is displayed as prefix to the position nucleotide. The nucleotide of iSNV is suffixed at the position of nucleotide. The non-synonymous substitution is presented as the following sequence, protein: dominant, residue location, and mutant. The dash sign indicates the stop codon. The distribution of the mutant percentage, that is the iSNV percentage, is shown as median with lower and upper hinges corresponding to the first and third quartiles on the panels. The statistical significance of difference between groups is calculated by the Mann-Whitney *U* test. Asterisks indicate the statistical significance (**p* <0.05, ***p* <0.01, ****p* <0.005). IFN, interferon; IFN-NR, IFN-non-responders; IFN-RS, IFN-responders; iSNVs, intra-host single-nucleotide variants; TN-eSC, treatment-naïve HBeAg seroconverters; TN-No-eSC, treatment-naïve non-seroconverters.

Discussion

In this study, we demonstrate the utility of a cost-effective amplicon-based deep sequencing approach in examining the evolution and quasispecies dynamics of nearly the whole HBV genome along with the course of HBeAg seroconversion, either spontaneous or IFN-induced. We found that the viral quasispecies diversity, measured by nucleotide diversity of HBV, was inversely correlated with serum viral load and ALT levels. The quasispecies diversity of HBV differed significantly between spontaneous HBeAg seroconverters and non-seroconverters and between IFN-treatment responders and non-responders. Of note, higher genetic diversity in treatment-naïve HBeAg seroconverters occurs at greater than 1 year before the occurrence of HBeAg seroconversion, suggesting that the patients with higher viral quasispecies diversity may have a higher chance to undergo spontaneous HBeAg seroconversion in the ensuing years. Also, HBeAg-positive CHB patients who remained at a low level of viral quasispecies diversity during IFN treatment were less likely to achieve HBeAg seroconversion as well as viral suppression following IFN treatment. Taken together, the findings indicate that viral quasispecies diversity can serve as a viral biomarker in the prediction of spontaneous HBeAg seroconversion and the response to IFN treatment.

A previous study has shown that higher genetic diversity of HBV core gene was associated with HBeAg seroconversion.[17] Consistently, our results demonstrated that the quasispecies diversity of HBV was higher in treatment-naïve HBeAg seroconverters. However, there was no significant difference of HBV quasispecies diversity between IFN-induced HBeAg seroconverters and non-seroconverters. Instead, the significantly differential HBV quasispecies diversity was observed between IFN-treatment responders (HBeAg seroconversion with viral load <20,000 IU/ml) and IFN-treatment non-responders. This result points out the potentially different impact of selection pressure on HBV genome between the spontaneous and IFN-induced HBeAg seroconverters. Although

the underlying mechanisms remain unclear, 1 plausible explanation is that the sole HBeAg seroconversion induced by IFN treatment does not necessarily result in good viral suppression, which is reflected by the increased viral quasispecies diversity. A prior study reported that IFN-induced HBeAg seroconversion is quite durable,[39] but there might exist a biological difference between spontaneous and IFN-induced HBeAg seroconversion.

The negative correlation between HBV genetic diversity and viral load has been found in several prior studies.[15,16] One plausible explanation for this observation is that the increased immune selection pressure will suppress viral replication and promote the escape mutations, which will consequently increase the viral genetic diversity. This speculation also suggests that even the selection pressure is strong enough to suppress viral replication, it cannot achieve complete inhibition of ongoing viral replication, which allows for the emergence of viral escape mutations and diversification. However, we cannot exclude the possibility that the increased viral genetic diversity may result in the reduction of viral population due to loss of viral fitness. Future studies are required to solve this issue by analysing the host antiviral immunity and viral quasispecies evolution contemporaneously.

To examine whether T cell immunity imposes selection pressure on viral quasispecies, we measured viral genetic diversity within and outside the predicted T cell epitopes. However, we found no significant difference of viral nucleotide diversity between within and outside the predicted T cell epitopes. Nevertheless, the interpretation of these results should be cautious because prediction of T cell epitopes based on common HLA alleles may be not accurate, and may have some intrinsic bias. Therefore, future studies with known HLAs and more accurate information of T cell epitopes are required to further investigate this issue in the future.

The NGS approach also allows for effective detection of viral variants at as low frequency as 1%.[40] Therefore, by NGS, we were able to discover iSNVs, defined as >5% frequency of the minor allele. Our results showed that the normalised number of iSNV was higher in spontaneous HBeAg seroconverters and IFN-treatment responders. We further identified several non-synonymous iSNVs in the non-overlapping region that were prevalent and had differential major allele frequency between patients with distinct clinical outcomes. Intriguingly, almost all these iSNVs were located within the class I HLA epitopes. Given that antiviral T cell immunity is critical for the control of HBV infection, it is not surprising that the iSNVs with non-synonymous substitutions might exist within the immune epitopes and are associated with the escape from host immunity. Although this study did not obtain the patients' HLA information, our findings warrant further investigations to determine the relationship between the iSNVs and T cell epitopes in future studies.

Interestingly, we also observed that the variation patterns of genetic distance in both treatment-naïve HBeAg non-seroconverters and IFN-treatment responders were correlated with that within HBV-B, and that between HBV genotypes (HBV-A, -B, -C, and -D). At the first glance, the finding seems surprisingly counterintuitive given that both spontaneous HBeAg seroconverters and IFN-treatment responders exhibited similarly higher HBV quasispecies diversity. However, we reasoned that TN-No-eSC tend to have high viraemia, which confers a higher risk for viral transmission. Therefore, the HBV genetic variations in the TN-No-eSC patients are theoretically more likely to contribute to the diversification within a population, and evolve in a pattern similar to that within genotype B and to that between genotypes. In addition, the individuals in the TN-No-eSC group are the CHB patients who require antiviral treatment. As expected, the pretreatment viral genetic diversity of the IFN-treated patients was as low as that of the TN-No-eSC group. However, IFN treatment enhanced the levels of HBV genetic diversity, particularly in the IFN-treatment responders. The correlation of the genetic variation patterns with that within genotype B and with that between genotypes strongly suggests that the driving force leading to the diversification of HBV quasispecies within an individual with CHB could be amplified by IFN treatment in the responders. It is likely that although the immune pressure in TN-No-eSC patients fails to effectively induce HBeAg seroconversion, but can drive viral evolution, eventually leading to the HBV diversification within and between genotypes.

In conclusion, the detailed analysis of HBV quasispecies dynamics during HBeAg seroconversion uncovers the viral evolution and adaptation under host immune selective pressure. HBV quasispecies diversity may serve as a novel biomarker to predict spontaneous HBeAg seroconversion.

Abbreviations

ALT, alanine aminotransferase; AUC, area under curve; BCP, basal core promoter; C, core; CHB, chronic hepatitis B; dN, nonsynonymous substitution rate; dS, synonymous substitution rate; EOT, end of treatment; IFN, interferon; IFN-eSC, IFN-treated HBeAg seroconverters; IFN-No-eSC, IFN-treated HBeAg non-seroconverters; IFN-NR, IFN-non-responders; IFN-RS, IFN-responders; iSNVs, intra-host single-nucleotide variants; NGS, next-generation sequencing; ORFs, open reading frames; P, polymerase; S, surface; TN-eSC, treatment-naïve HBeAg seroconverters; TN-No-eSC, treatment-naïve non-seroconverters.

Financial support

This work was supported by the grants from the Ministry of Science and Technology, Taiwan (MOST105-2314-B-002-178-MY3 to H.C.Y. and MOST109-2314-B-002-069-MY3 to J.H.K.); the Ministry of Education, Taiwan (CTCN-104-01 to I.J.L.); National Taiwan University Hospital (NTUH105-003046 to H.C.Y.); National Taiwan University, Taiwan (NTU109L890301 to H.C.Y. and J.H.K.).

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Conceived and designed this study: H.C.Y., J.H.K. Designed and performed the experiments: S.R.L., T.Y.Y., I.J.L., H.R.C., Y.C.S., F.Y.W. Contributed to sample collection and clinical consultation: C.Y.P., C.Y.D., T.H.S., T.C.T., C.J.L., P.J.C., D.S.C., H.C.Y., J.H.K. Analysed and interpretated data: S.R.L., T.Y.Y., Y.Y.L., H.Y.W., H.C.Y. Wrote the manuscript with critical feedback from all authors: S.R.L., T.Y.Y., H.C.Y., J.H.K.

Data availability statement

The raw FASTQ files are available in the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under the accession number PRJNA556730.

Acknowledgements

We thank Dr Pei-Lung Chen from the Laboratory of Molecular Genetic Diagnostics, Department of Medical Genetics at the National Taiwan University Hospital for technical support. We would like to acknowledge the service provided by Dr Huei-Ying Li and the Medical Microbiota Center of the First Core Laboratory, National Taiwan University College of Medicine. We also thank Yao-Cheng Lin and Po-Xing Zheng for helpful discussion and computational support from the Bioinformatics and Computational Genomics Lab, Agricultural Biotechnology Research Center, Academia Sinica, Taiwan.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhepr.2021.100254.

References

Author names in bold designate shared co-first authorship

[1] Kao JH. Diagnosis of hepatitis B virus infection through serological and virological markers. Expert Rev Gastroenterol Hepatol 2008;2:553–562.
[2] Coffin CS, Zhou K, Terrault NA. New and old biomarkers for diagnosis and management of chronic hepatitis B virus infection. Gastroenterology 2019;156. 355–68.e3.
[3] Yang HC, Shih YF, Liu CJ. Viral factors affecting the clinical outcomes of chronic hepatitis B. J Infect Dis 2017;216:S757–S764.
[4] Seeger C, Mason WS. Molecular biology of hepatitis B virus infection. Virology 2015;479–480:672–686.
[5] Nowak MA, Bangham CR. Population dynamics of immune responses to persistent viruses. Science 1996;272:74–79.
[6] Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 2006;439:344–348.
[7] Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. Microbiol Mol Biol Rev 2012;76:159–216.
[8] European Association for the Study of the Liver. EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. J Hepatol 2017;67:380–398.
[9] Terrault NA, Lok ASF, McMahon BJ, Chang KM, Hwang JP, Jonas MM, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. Hepatology 2018;67:1560–1599.
[10] Kao JH. HBeAg-positive chronic hepatitis B: why do I treat my patients with pegylated interferon? Liver Int 2014;34(Suppl 1):112–119.
[11] Buster EH, Flink HJ, Cakaloglu Y, Simon K, Trojan J, Tabak F, et al. Sustained HBeAg and HBsAg loss after long-term follow-up of HBeAg-positive patients treated with peginterferon alpha-2b. Gastroenterology 2008;135:459–467.
[12] van Zonneveld M, Honkoop P, Hansen BE, Niesters HG, Darwish Murad S, de Man RA, et al. Long-term follow-up of alpha-interferon treatment of patients with chronic hepatitis B. Hepatology 2004;39:804–810.
[13] Yuen MF, Sablon E, Yuan HJ, Hui CK, Wong DK, Doutreloigne J, et al. Relationship between the development of precore and core promoter mutations and hepatitis B e antigen seroconversion in patients with chronic hepatitis B virus. J Infect Dis 2002;186:1335–1338.
[14] Yang HC, Chen CL, Shen YC, Peng CY, Liu CJ, Tseng TC, et al. Distinct evolution and predictive value of hepatitis B virus precore and basal core promoter mutations in interferon-induced hepatitis B e antigen seroconversion. Hepatology 2013;57:934–943.
[15] Wang HY, Chien MH, Huang HP, Chang HC, Wu CC, Chen PJ, et al. Distinct hepatitis B virus dynamics in the immunotolerant and early immunoclearance phases. J Virol 2010;84:3454–3463.
[16] Cheng Y, Guindon S, Rodrigo A, Wee LY, Inoue M, Thompson AJ, et al. Cumulative viral evolutionary changes in chronic hepatitis B virus infection precedes hepatitis B e antigen seroconversion. Gut 2013;62:1347–1355.
[17] Lim SG, Cheng Y, Guindon S, Seet BL, Lee LY, Hu P, et al. Viral quasi-species evolution during hepatitis Be antigen seroconversion. Gastroenterology 2007;133:951–958.
[18] Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 2008;9:387–402.
[19] Rodriguez-Frias F, Buti M, Tabernero D, Homs M. Quasispecies structure, cornerstone of hepatitis B virus infection: mass sequencing approach. World J Gastroenterol 2013;19:6995–7023.
[20] McNaughton AL, D'Arienzo V, Ansari MA, Lumley SF, Littlejohn M, Revill P, et al. Insights from deep sequencing of the HBV genome – unique, tiny, and misunderstood. Gastroenterology 2019;156:384–399.
[21] Chevaliez S, Rodriguez C, Pawlotsky JM. New virologic tools for management of chronic hepatitis B and C. Gastroenterology 2012;142:1303–13013.e1.
[22] Liu CJ, Lai MY, Chao YC, Liao LY, Yang SS, Hsiao TJ, et al. Interferon alpha-2b with and without ribavirin in the treatment of hepatitis B e antigen-positive chronic hepatitis B: a randomized study. Hepatology 2006;43:742–749.
[23] Gregori J, Perales C, Rodriguez-Frias F, Esteban JI, Quer J, Domingo E. Viral quasispecies complexity measures. Virology 2016;493:227–237.
[24] Wang YC, Yang SS, Su CW, Wang YJ, Lee KC, Huo TI, et al. Predictors of response to pegylated interferon in chronic hepatitis B: a real-world hospital-based analysis. Sci Rep 2016;6:29605.
[25] Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999;50:213–219.
[26] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res 2019;47:D339–D343.
[27] Parker KC, Bednarek MA, Coligan JE. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J Immunol 1994;152:163–175.
[28] Doytchinova IA, Guan P, Flower DR. EpiJen: a server for multistep T cell epitope prediction. BMC Bioinformatics 2006;7:131.
[29] **Andersen KG**, **Shapiro BJ**, **Matranga CB**, Sealfon R, Lin AE, Moses LM, et al. Clinical sequencing uncovers origins and evolution of lassa virus. Cell 2015;162:738–750.
[30] Hayer J, Jadeau F, Deleage G, Kay A, Zoulim F, Combet C. HBVdb: a knowledge database for hepatitis B virus. Nucleic Acids Res 2013;41:D566–D570.
[31] van der Burg SH, Visseren MJ, Brandt RM, Kast WM, Melief CJ. Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability. J Immunol 1996;156:3308–3314.
[32] Depla E, Van der Aa A, Livingston BD, Crimi C, Allosery K, De Brabandere V, et al. Rational design of a multiepitope vaccine encoding T-lymphocyte epitopes for treatment of chronic hepatitis B virus infections. J Virol 2008;82:435–450.
[33] Desmond CP, Gaudieri S, James IR, Pfafferott K, Chopra A, Lau GK, et al. Viral adaptation to host immune responses occurs in chronic hepatitis B virus (HBV) infection, and adaptation is greatest in HBV e antigen-negative disease. J Virol 2012;86:1181–1192.
[34] Malmassari S, Lone YC, Zhang M, Transy C, Michel ML. In vivo hierarchy of immunodominant and subdominant HLA-A*0201-restricted T-cell epitopes of HBx antigen of hepatitis B virus. Microbes Infect 2005;7:626–634.
[35] King TH, Kemmler CB, Guo Z, Mann D, Lu Y, Coeshott C, et al. A whole recombinant yeast-based therapeutic vaccine elicits HBV X, S and core specific T cells in mice and activates human T cells recognizing epitopes linked to viral clearance. PLoS One 2014;9:e101904.
[36] **Chung MK**, **Yoon H**, Min SS, Lee HG, Kim YJ, Lee TG, et al. Induction of cytotoxic T lymphocytes with peptides in vitro: identification of candidate T-cell epitopes in hepatitis B virus X antigen. J Immunother 1999;22:279–287.
[37] Kefalakes H, Jochum C, Hilgard G, Kahraman A, Bohrer AM, El Hindy N, et al. Decades after recovery from hepatitis B and HBsAg clearance the CD8+ T cell response against HBV core is nearly undetectable. J Hepatol 2015;63:13–19.
[38] Lumley S, Noble H, Hadley MJ, Callow L, Malik A, Chua YY, et al. Hepitopes: a live interactive database of HLA class I epitopes in hepatitis B virus. Wellcome Open Res 2016;1:9.
[39] Lau DT, Everhart J, Kleiner DE, Park Y, Vergalla J, Schmid P, et al. Long-term follow-up of patients with chronic hepatitis B treated with interferon alfa. Gastroenterology 1997;113:1660–1667.
[40] Van den Hoecke S, Verhelst J, Vuylsteke M, Saelens X. Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. BMC Genomics 2015;16:79.