ORIGINAL RESEARCH

Quasispecies characteristic in "a" determinant region is a potential predictor for the risk of immunoprophylaxis failure of mother-to-child-transmission of sub-genotype C2 hepatitis B virus: a prospective nested case–control study

Yiwei Xiao [ID],[1] Kuixia Sun,[1,2] Zhongping Duan,[3] Zhixiu Liu,[1] Yi Li,[1] Ling Yan,[1] Yarong Song,[1] Huaibin Zou,[3] Hui Zhuang,[1] Jie Wang,[1] Jie Li[1]

[1]Department of Microbiology & Infectious Disease Center, School of Basic Medical Sciences, Peking University Health Science Center, Beijing, 100083, P.R. China
[2]Department of Clinical Laboratory, Peking University First Hospital, Beijing, 100034, P.R. China
[3]Artificial Liver Center, Beijing Youan Hospital, Capital Medical University, Beijing, 100054, P.R. China

Correspondence to
Associate Professor Jie Li, Associate Professor Jie Wang and Professor Hui Zhuang, Department of Microbiology & Infectious Disease Center, School of Basic Medical Sciences, Peking University Health Science Center, Beijing, 100083, P.R. China; jieli@bjmu.edu.cn, wangjie2015@bjmu.edu.cn, zhuangbmu@126.com

YX and KS contributed equally.

Check for updates

ABSTRACT

Objective This study was performed to explore the correlation between the characteristics of hepatitis B virus (HBV) quasispecies in HBV-infected pregnant women and the risk of immunoprophylaxis failure for their infants.

Design In this prospective nested case–control study, the characteristics of HBV quasispecies in mothers whose infants were immunoprophylaxis success (control group) and those whose infants were immunoprophylaxis failure (case group) were analysed by the clone-based sequencing of full-length HBV genome and next-generation sequencing (NGS) of "a" determinant region, and were compared between the two groups.

Results The quasispecies characteristics including mutant frequency, Shannon entropy and mean genetic distance at amino acid level of "a" determinant region were significantly lower in case group than that in control group, using the full-length HBV genome clone-based sequencing assay. These results were confirmed by NGS assay. Notably, we discovered that the differences were also significant at nucleotide level by NGS assay. Furthermore, the risk of immunoprophylaxis failure could be predicted by analysing the three HBV quasispecies characteristics either at nucleotide level or at amino acid level of "a" determinant region, and the corresponding predictive values were tentatively set up.

Conclusions HBV quasispecies with a more complex mutant spectrum in "a" determinant region might be more vulnerable to extinct through mother-to-child-transmission (MTCT). More importantly, analysing HBV quasispecies characteristics in pregnant women with high HBV DNA load might be helpful to predict the high-risk population of immunoprophylaxis failure, and consequently provide accurate intervention against MTCT of HBV.

Significance of this study

What is already known about this subject?
► Positive hepatitis B e antigen and high hepatitis B virus (HBV) DNA load are associated with the high risk of mother-to-child-transmission (MTCT) of HBV, even though the active and passive immunisations were taken to their infants timely.
► To reduce the risk of HBV MTCT, antiviral therapy is recommended in the HBV-infected pregnant women with high HBV DNA load in the third trimester of pregnancy.
► The efficacy and safety of antiviral therapy in preventing HBV MTCT remain controversial.
► So far, the indicator used for predicting the risk of immunoprophylaxis failure for their infants has not been found in the pregnancy women with high HBV DNA load.

What are the new findings?
► HBV evolution patterns were similar between mothers whose infants were immunoprophylaxis success (control group) and those whose infants were immunoprophylaxis failure (case group).
► The quasispecies characteristics including mutant frequency, Shannon entropy and mean genetic distance in "a" determinant region of HBV genome were significantly lower in case group than that of control group.
► These HBV quasispecies characteristics are the potential markers for predicting the risk of immunoprophylaxis failure for their infants in the pregnant women with high HBV DNA load.

INTRODUCTION

Chronic hepatitis B virus (HBV) infection remains a serious threat to public health and is associated with many liver diseases such as chronic hepatitis, cirrhosis and liver cancer.[1] There are 257 million persons chronically infected with HBV in the world, and 887 000 people died of HBV-related end-stage liver disease including cirrhosis, liver failure and liver cancer annually.[1 2]

Mother-to-child transmission (MTCT) is one of the predominant routes of HBV dissemination worldwide, especially in the areas with high prevalence of HBV. Fortunately, over 90% infants born to hepatitis B surface antigen (HBsAg)-positive mothers

How might it impact on clinical practice in the foreseeable future?

► Analysing HBV quasispecies characteristics, including mutant frequency, Shannon entropy and mean genetic distance in "a" determinant region, may be helpful to predict the high-risk populations of immunoprophylaxis failure for their infants in pregnant women with high HBV DNA load, and consequently provide accurate medication to improve the safety and effectiveness in preventing MTCT of HBV.

can be protected from MTCT, due to the universal implementation of hepatitis B vaccine strategy in newborns.[3] In addition, the chronically HBV-infected children under 5-year old reduced from 4.7% to 1.3%,[2] which is attributed to the continuous improvements of vaccine immunisation strategies, including three-dose coverage, timely birth dose, additional dose of HBV immunoglobulin (HBIG) for neonates born to HBsAg-positive mothers within 24 hours post-delivery, and post-vaccination serological testing (PVST), etc. In 2016, the World Health Assembly endorsed the Global Health Sector Strategy on viral hepatitis, with elimination of viral hepatitis as a public health threat based on reducing of incidence of chronic hepatitis B by 90% and mortality related to HBV by 65% (from the 2015 baseline), as well as HBsAg prevalence to 0.1% among children under 5 years by 2030.[2 4] To achieve these goals, MTCT of HBV is still a challenging issue in the world because MTCT is responsible for about 50% of HBV new infection so far.[5] Accordingly, preventing HBV infection at birth and in the first year of life should be a high priority in the driving towards elimination of hepatitis B.

It has been reported that the risks of HBV MTCT are associated with maternal hepatitis B e antigen (HBeAg) status and serum HBV DNA load.[6–8] To reduce the risk of HBV MTCT, antiviral therapy is recommended for HBV-infected pregnant women with high HBV DNA load during the third trimester because some studies have shown its benefits in preventing MTCT of HBV.[9 10] However, it has also been reported that the additional antiviral therapy of tenofovir disoproxil fumarate for pregnant women with high HBV DNA load does not result in a significantly lower rate of HBV MTCT compared with immunoprophylaxis alone.[11] Several reports have also shown that the potential severe harms of the antiviral agent to mothers or neonates should be a concern based on the limited fetal safety data of antiviral agent; thus, the antiviral intervention strategy for preventing MTCT of HBV remains controversial.[10 12 13] Moreover, since only 5%–10% of infants born to high HBV DNA load mothers without maternal antiviral therapies are immunoprophylaxis failure, it warranted to predict these mothers with high risk of immunoprophylaxis failure for their infants to reduce the medical expenses and the potential side effects of antiviral therapy.[6 8 14] For virological factors, uncovering the HBV virological mechanisms related to immunoprophylaxis failure may help us to predict the high-risk population of HBV MTCT, so as to provide precision medication for HBV-infected pregnant women with high HBV DNA load. However, the exact virological mechanisms associated with immunoprophylaxis failure in infants are not well understood yet. Several studies have demonstrated that immunoprophylaxis failure is associated with mutations of HBV genome, such as the mutations occurring in "a" determinant region may result in immune escape, and contribute to immunoprophylaxis failure.[15 16] Meanwhile, several studies focus on the comparison of quasispecies

characteristics between immunoprophylaxis failure infants and their mothers to demonstrate the HBV evolution patterns in MTCT.[17–19] It has been suggested that the evolution patterns of HBV quasispecies between genotype B and C during vertical transmission were different, which may contribute to the distinct prognosis.[19] However, the correlation between the maternal HBV quasispecies characteristics and the immunoprophylaxis outcomes of their infants is still inconclusive.

In this study, using a large prospective cohort of mother–infant pairs with positive maternal HBsAg,[8 20] a prospective nested case–control study was performed to analyse the characteristics of HBV quasispecies between the mothers whose infants were immunoprophylaxis failure and those whose infants were immunoprophylaxis success, and to explore the correlation between the maternal HBV quasispecies characteristics and the immunoprophylaxis outcomes of their infants.

MATERIALS AND METHODS
Subject
As previously reported, pregnant women recruited from county-level community maternal and child healthcare centres from August 2009 to June 2011 in Jiangsu and Henan provinces were routinely tested for HBsAg during the first prenatal visit by 12 weeks of gestation.[9] HBsAg-positive pregnant women were recommended to join the study at the time of prenatal visits. All pregnant women underwent comprehensive evaluation for the need of antiviral therapy by clinicians, based on age, family history, alanine aminotransferase (ALT), HBeAg status, viral load, disease severity and personal willingness. Those with severe hepatitis, fibrosis and cirrhosis or at high risk of disease progression, and those who commenced antiviral therapy during or before pregnancy were excluded. All the enrolled pregnant women were negative for hepatitis A virus, hepatitis C virus, hepatitis D virus, hepatitis E virus and HIV.

In this study, 1448 mother–infant pairs with positive maternal HBsAg were enrolled. In all, 1446 infants were received three doses of recombinant yeast-derived hepatitis B vaccine (10 μg/0.5 mL; Shenzhen Kangtai Biological Products Co., Shenzhen, China or Dalian Hissen Biopharm, Dalian, China) intramuscularly at birth (within 12 hours), 1 month and 6 months, combined with one dosage of HBIG (100 or 200 IU/bottle; Hualan Biological Engineering, Xinxiang, China) within 12 hours of birth. In total, 1177 infants were returned for PVST at 7 months and were then followed up to 12 months.

Totally, 20 infants were immunoprophylaxis failure after the full course of vaccination, 15 of 20 infants were infected with genotype C2 HBV. All of the immunoprophylaxis failure infants were born to mothers with high viral loads (median, 8.38; range: 7.82–9.22 log IU/mL), the immunoprophylaxis failure rate was 6.1% (20/324).[9] Since only five infants were infected with other genotypes HBV, these five infants were excluded to reduce the influence from the high heterogeneity of different genotypes. Immunoprophylaxis failure was defined as positive HBsAg and antibody to HBsAg (anti-HBs) levels <10 mIU/mL on follow-up serological testing at 7 and 12 months after the first dose of HBV vaccine.[21 22] Based on age, HBeAg status, HBV genotype, serum HBV DNA level, ALT level and total bilirubin level, 30 of 106 eligible mothers were randomly selected in the control group matched to 15 mothers in the case group (figure 1). Since more HBV full-length genome clones would be needed if more controls per case were designed in this study and the power gained by increasing to more than two controls per case is little for statistical analyses,[23 24] the ratio of 1:2 was selected to
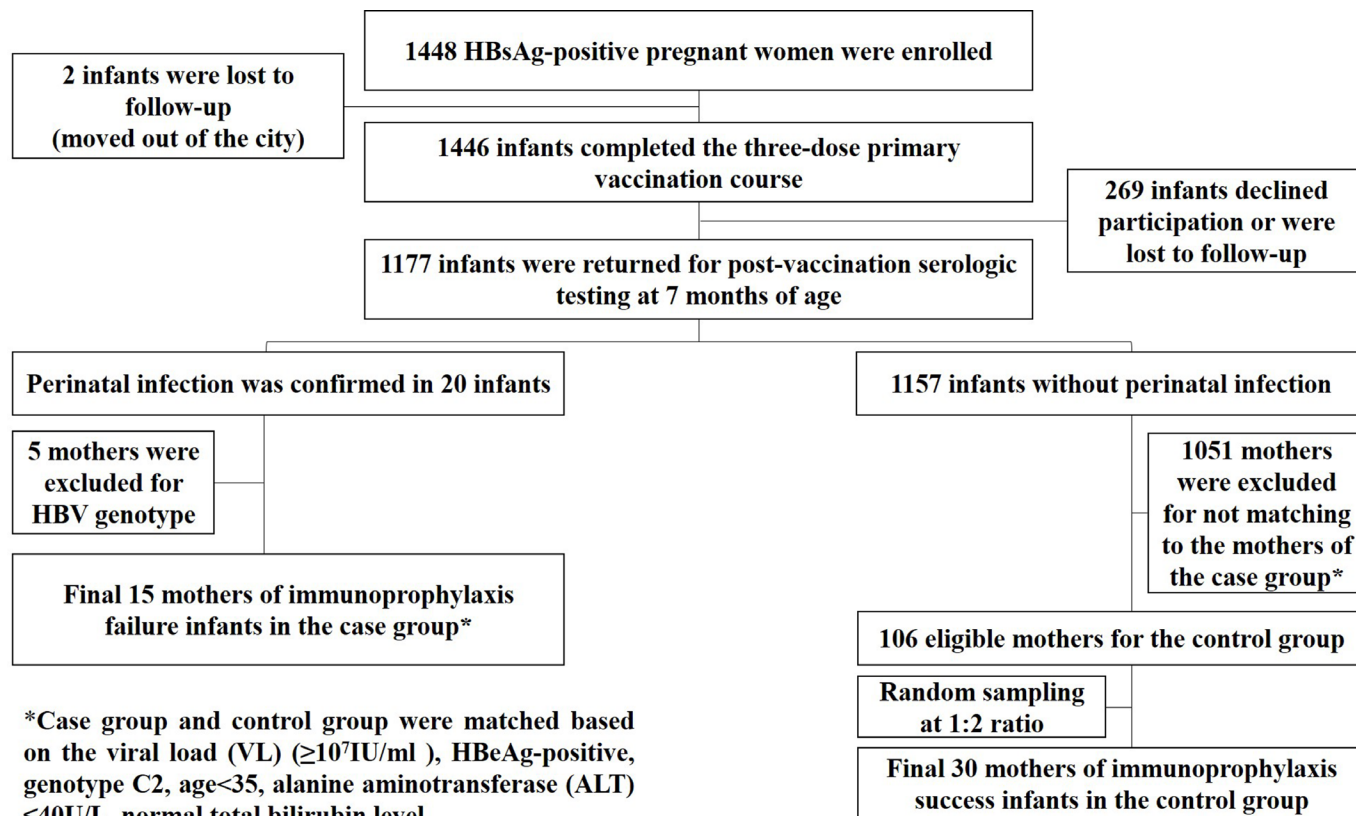
Figure 1 Flowchart of case and control group selection. HBV, hepatitis B virus; HBsAg, hepatitis B surface antigen.

balance the cost-effectiveness of the subsequent full-length HBV genome clone-based sequencing assays and sequences analyses in this study.
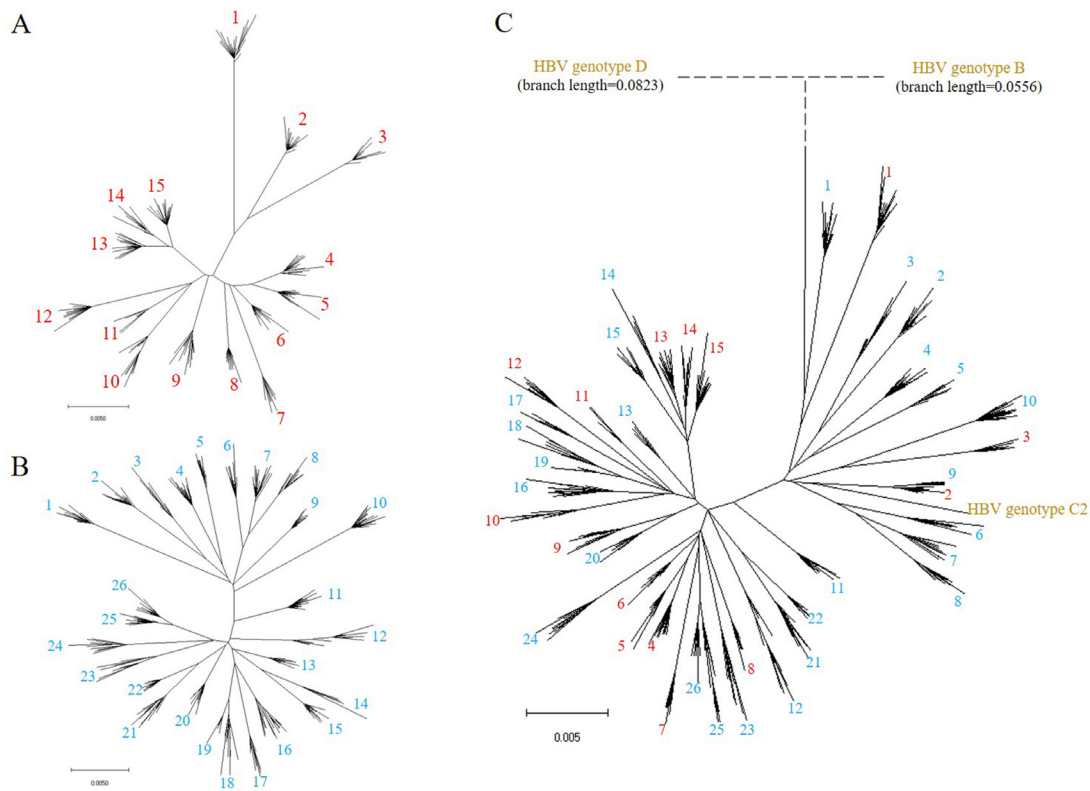
Baseline characteristics of mothers and their infants are described in table 1. No significant differences were detected in the baseline characteristics of mothers and their infants (Table 1 and online supplementary table S1).

Amplification, cloning and sequencing

HBV genomes were extracted from 200 μL serum samples using QIAamp DNA blood mini kit (Qiagen, Hilden, Germany). The full-length HBV genomes were amplified by PCR as described by Gunther et al,[25] followed by clone-based sequencing. Of 30 mothers in control group, four mothers failed full-length clone amplification. After four mothers excluded also, no significant differences were detected in the baseline characteristics of mothers and their infants either. A total of 208 clones in case group and 344 clones in control group were sequenced using seven sequencing primers (online supplementary table S2). The detail information is described in supplementary materials.

Sequence analyses

Sequence segments were assembled to full-length HBV genome using the Contig-Express software (New York, USA) and Codon Code Aligner software (Centerville, MA, USA). All mutations were checked manually. Sequence alignment and viral quasispecies analyses were performed by MEGAX software.[26] The following regions were aligned with the reference sequences (genotype C2, GenBank accession no. AB014378 and genotype C2, GenBank accession no. AB048705) using Clustal X (V.2.0) software: "a" determinant, nt 524–595; HBcAg region (HBcAg), nt 1901–2452; polymerase region (P), nt 2307–1623; NTCP-binding domain (NTCP-BD), nt 2851–3072; PreC region (PreC), nt 1814–1900; PreS1 region (PreS1), nt 2848–3204; PreS2 region (PreS2), nt 3205–154; RT region (RT), nt

Table 1 Baseline characteristics of mothers and their infants

Characteristics	Case group	Control group	P value
Mothers			
No. of Mothers	15	30	
Age (years)	23.4±4.067*	23.33±2.339	0.954
Alanine aminotransferase (U/L)	17.75±7.768	18.86±7.392	0.650
Hepatitis B surface antigen (IU/mL)	35766.1±18 249	30243.1±18 035	0.345
Hepatitis B e antigen (S/CO)	1282.7±293.3	1259.3±314.7	0.807
Hepatitis B virus DNA-$\log_{10}$ (IU/mL)	8.41±0.398	8.24±0.414	0.207
Infants			
Gender of infant			
Male	8 (53.3%)	16 (53.3%)	1.000
Female	7 (46.7%)	14 (46.7%)	
Method of delivery			
Caesarean	7 (46.7%)	19 (63.3%)	0.286
Natural	8 (53.3%)	11 (36.7%)	
Type of feeding			
Breast fed	3 (20.0%)	1 (3.3%)	0.127
Bottle fed	9 (60.0%)	25 (83.3%)	
Mixed	3 (20.0%)	4 (13.3%)	
Weight (kg)	3.52±0.3326	3.32±0.3042	0.061

P values were calculated by student's t-test and Fisher's exact test.
*Values were means±SD.

Figure 2 The phylogenetic analysis of the HBV genome sequences in mothers between the case group and control group. The maximum likelihood phylogenetic trees reconstructed for the cloned HBV genome sequences in mothers of case group (A), control group (B) and two groups (C). The red numbers represented the mothers of case group, the blue numbers represented the mothers of control group and the yellow represented the reference sequence of genotype B (GenBank accession No. LC057377), C2 (GenBank accession No. AB014378) or D HBV (GenBank accession No. LT718449). HBV, hepatitis B virus.

130–1161; HBsAg region (HBsAg), nt 155–835; HBx protein region (HBx), nt 1374–1838; enhancer I (EN I), nt 957–1361; enhancer II (EN II), nt 1685–1773; basic core promoter (BCP), nt 1742–1849; core promoter (CP), nt 1613–1849; S promoter I (SP I), 2718–2808; S promoter II (SP II), nt 2983–3210; X promoter (XP), nt 1171–1361.[27]

Viral quasispecies characterisation was evaluated by three parameters: mutation frequency, complexity (Shannon entropy) and diversity (genetic distance),[28] all calculated values are shown in online supplementary 2. All the gaps caused by deletion mutation in the aligned sequences were removed, and the protein sequences were translated from the processed nucleotide sequences. Subsequently, the amino acid sequences were exported to Excel, and the sequences behind the first stop codon were removed. These analyses are described in detail in supplementary materials. Phylogenetic analyses were performed by maximum likelihood tree constructed by general time reversible (GTR) model given in the MEGAX software. The reliability of pairwise comparison and phylogenetic tree analyses were assessed by 1000 replicates bootstrapping. Trees were drawn and rearranged by MEGAX and Evolview.

Next-generation sequencing
The "a" determinant region of HBV was amplified by the PCR using PrimerSTAR MAX (TAKARA) under the following PCR conditions: 98°C 2 min; 98°C 10 s, 55°C 5 s and 72°C 5 s for 35 cycles; 72°C 5 min. The primers were Af2 (5′-TCCAGGAACA TCAACTACCAG-3′; nt 484–505) and Ar2 (5′-AGGGTTCA AATGTATACCCAA-3′; nt 840–820). The 150 bp paired-end libraries were sequenced by Illumina Hiseq X. The details of

NGS workflow are shown in online supplementary figure S1. The target reads of "a" determinant region from libraries are shown in online supplementary table S3, and there was no significant difference (p=0.529) between the two groups. Three quasispecies characteristics were calculated by self-written perl scripts (provided by Shanghai OE Biotechnology Co.) at nucleotide and amino acid level, respectively. All calculated values are shown in online supplementary 2.

Serological assays
Serum HBsAg, HBeAg, anti-HBs, antibody to HBeAg (anti-HBe) and antibody to HBcAg (anti-HBc) were tested by Abbott chemiluminescent microparticle immunoassay (Abbott Diagnostic, Chicago, IL, USA) using Abbott i2000 system.[29] HBV DNA load was measured by Abbott real-time HBV DNA assay (Abbott Molecular, IL, USA) using Abbott m2000 system. Serum HBV DNA load was reported as 'not detected' or '<1.18 log IU/mL' when its level was below the lower detection limit of 1.18 log IU/mL.[29] HBV genotyping was performed by a method based on nested PCR as described previously.[30]

Statistical analyses
Categorical variables were expressed as % (m/n) and examined by χ^2/Fisher's exact test. Normal distributions data were expressed as mean±SD and compared by the student's t-test. Non-normal distributions data were expressed as median and IQR or (range) and compared by Mann-Whitney U test. Non-parametric Spearman and binary logistic regression

Table 2 Comparing the quasispecies characteristics in "a" determinant region at amino acid and nucleotide level based on clone-based sequencing assay between case group and control group

Region of variation (level)	Case group (n=15) Median (IQR) (range)	Control group (n=26) Median (IQR) (range)	P value
Mutant frequency (nt)	0.0011 (0.0008–0.0016) (0–0.0038)	0.0024 (0.0008–0.0042) (0–0.0135)	0.1345
Mutant frequency (aa)	0 (0–0.0038) (0–0.0066)	0.0040 (0.0017–0.0117) (0–0.0417)	**0.0094**
Shannon entropy (nt)	0.1154 (0.07896–0.1588) (0–0.2739)	0.1631 (0.07423–0.3664) (0–0.6847)	0.2165
Shannon entropy (aa)	0 (0–0.1391) (0–0.2070)	0.1341 (0.0557–0.3088) (0–0.6847)	**0.0121**
Genetic distance (nt)	0.0023 (0.0017–0.0033) (0–0.0060)	0.0036 (0.0016–0.0081) (0–0.0218)	0.2115
Genetic distance (aa)	0 (0–0.0093) (0–0.0187)	0.0084 (0.0037–0.0226) (0–0.0653)	**0.0152**

Pvalues were calculated by Mann-Whitney test.
aa, amino acid; nt, nucleotide.

models were used when appropriate. Logistic regression was performed to develop the predictive model. All P values were two-tailed and a P value<0.05 was considered significant. Statistical analyses were analysed using SPSS software, V.25.0 (Chicago, IL, USA) and receiver operating characteristic (ROC) analyses were performed by MedCalc 18.2.1 software (MedCalc software, Mariakerke, Belgium).

RESULTS
No significant differences on the sequences of full-length HBV genomes were found in the mothers between case group and control group

To compare the HBV genome sequences between the cases and controls, the full-length sequences of 3.2 kb HBV genome were analysed by PCR-based clone sequencing assays. The phylogenetic analysis result revealed that, both in the cases and controls, the cloned HBV genome sequences from the same mother were clustered together, and the branches of different mothers were clustered in different clades, which indicated that HBV genome in each mother were derived from different ancestor sequences (figure 2A, B). Consistently, as shown in figure 2C, such a phenomenon was also present when two groups combined together, which indicated a close evolutionary relationship of HBV genome in the same mother. Furthermore, the overall branch lengths in two groups were not obviously different, and there was no any clustering sign of the cloned HBV genomes between two groups, suggesting that the sequences of HBV genomes in case group were similar to those from control group. Moreover, all clones from two groups clustered together with the reference sequence of genotypes C2 HBV (GenBank accession no. AB014378), whereas the reference sequences of genotypes B (GenBank accession no. LC057377) and D HBV (GenBank accession no. LT718449) were scattered throughout the tree, suggesting that all cloned HBV genomes were genotype C2 (figure 2C). The detail branch lengths and bootstraps of the phylogenetic trees are shown in online supplementary figure S2.

The mutant frequency, Shannon entropy and genetic distance of HBV quasispecies at amino acid level in "a" determinant region were lower in the mothers of case group than that of control group in clone-based sequencing

The HBV quasispecies characteristics were calculated and compared between the mothers of case group and control group. As shown in table 2, the mutant frequency (0 (0–0.0038) vs 0.0040 (0.0017–0.0117), p=0.0094), Shannon entropy (0 (0–0.1391) vs 0.1341 (0.0557–0.3088), p=0.0121) and genetic distance (0 (0–0.0093) vs 0.0084 (0.0037–0.0226), p=0.0152) at amino acid level of "a" determinant region in the mothers of the case group were lower than that of the control group. However, there were no significant differences of the mutant frequency, Shannon entropy and genetic distance at nucleotide level of HBV "a" determinant region between the mothers of the case group and the control group (table 2). Meanwhile, no differences of these three quasispecies characteristics at amino acid level of other HBV coding regions, and at nucleotide level of other HBV coding and non-coding regions including PreC, PreS1, PreS2, HBsAg, ENII, SPI and SPII regions, were found between the mothers of the two groups (online supplementary tables S4–S6).

Furthermore, as shown in online supplementary table S7, the number of non-synonymous substitutions per non-synonymous site (dN) of "a" determinant region in the mothers of the case group were lower than that of the control group (0 (0–0.0039) vs 0.0035 (0.0015–0.0093), p=0.0179). However, there were also no differences of dN of other regions and the number of synonymous substitutions per synonymous site (dS) of all regions between the two groups (online supplementary table S7). Moreover, there were no significant differences of the mutant frequency and Shannon entropy at each amino acid position of HBV "a" determinant region between two groups (online supplementary figures S3 and S4).

HBV quasispecies characteristics in "a" determinant region could be used to predict the risk of immunoprophylaxis failure

To confirm the differences of three quasispecies characteristics in "a" determinant region between two groups found by clone sequencing assay, the NGS assay of "a" determinant region in HBV genome was performed. As shown in table 3, through the NGS assay, the significant differences of three characteristics parameters between two groups were found in "a" determinant region at amino acid level, which confirmed the results of clone sequencing assay. Notably, significant differences were also found at nucleotide level.

Table 3 Comparing the quasispecies characteristics in "a" determinant region at both nucleotide level and amino acid level based on NGS assay between case group and control group

Region of variation (level)	Case group (n=15) Median (IQR) (range)	Control group (n=26) Median (IQR) (range)	P value
Mutant frequency (nt)	0 (0–0.0002) (0–0.0042)	0.0004 (0–0.0027) (0–0.0068)	0.0210
Mutant frequency (aa)	0 (0–0.0006) (0–0.0126)	0.0009 (0–0.0058) (0–0.0205)	0.0490
Shannon entropy (nt)	0 (0–0.0055) (0–0.046)	0.0105 (0–0.0454) (0–0.1031)	0.0180
Shannon entropy (aa)	0 (0–0.0055) (0–0.046)	0.009 (0–0.0419) (0–0.0709)	0.0460
Genetic distance (nt)	0 (0–0.014) (0–0.3026)	0.0258 (0–0.1919) (0–0.4927)	0.0210
Genetic distance (aa)	0 (0–0.014) (0–0.3026)	0.0224 (0–0.1381) (0–0.4927)	0.0490

P values were calculated by Mann-Whitney test.
aa, amino acid; NGS, next-generation sequencing; nt, nucleotide.

Furthermore, the correlations between the immunoprophylaxis failure and HBV quasispecies characteristics were analysed by Spearman correlation analyses. The result revealed that there were negative correlations between the immunoprophylaxis failure and HBV quasispecies characteristics including the mutant frequency (r=−0.311, p=0.048), Shannon entropy (r=−0.316, p=0.044) and genetic distance (r=−0.311, p=0.048) at amino acid level, as well as the mutant frequency (r=−0.364, p=0.019), Shannon entropy (r=−0.373, p=0.016) and genetic distance (r=−0.364, p=0.019) at nucleotide level of "a" determinant region, respectively.

Next, the potential predictive values of quasispecies characteristics and HBV DNA level for immunoprophylaxis failure were analysed by the ROC curves. As shown in figure 3, the area under the ROC curve (AUC) for immunoprophylaxis failure due to the HBV quasispecies including mutant frequency, Shannon entropy and genetic distance were 0.677 (95% CI: 0.513 to 0.814, p=0.0247), 0.679 (95% CI: 0.516 to 0.816, p=0.0215) and 0.677 (95% CI: 0.513 to 0.814, p=0.0247) at amino acid level, as well as were 0.713 (95% CI: 0.550 to 0.843, p=0.0076), 0.718 (95% CI: 0.556 to 0.847, p=0.0055) and 0.713 (95% CI: 0.550 to 0.843, p=0.0076) at nucleotide
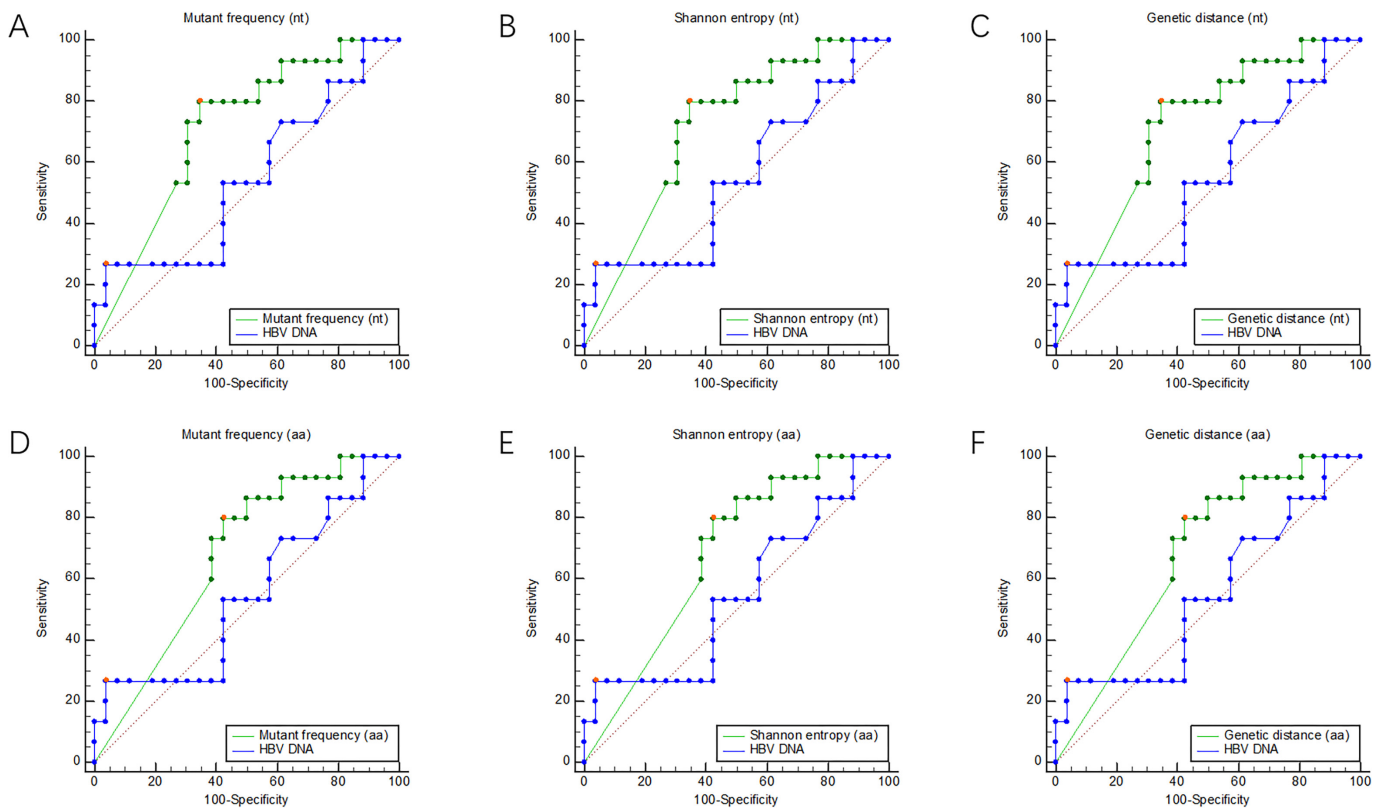
Figure 3 The predictive values of HBV quasispecies characteristics in "a" determinant for immunoprophylaxis failure using NGS assay were analysed by ROC curves. (A) The ROC curve of the mutant frequency in "a" determinant for predicting the risk of immunoprophylaxis failure at nucleotide level (AUC=0.713, 95% CI: 0.550 to 0.843, p=0.0076). (B) The ROC curve of the Shannon entropy in "a" determinant for predicting the risk of immunoprophylaxis failure at nucleotide level (AUC=0.718, 95% CI :0.556 to 0.847, p=0.0055). (C) The ROC curve of the quasispecies genetic distance in "a" determinant for predicting the risk of immunoprophylaxis failure at nucleotide level (AUC=0.713, 95% CI: 0.550 to 0.843, p=0.0076). (D) The ROC curve of the mutant frequency in "a" determinant for predicting the risk of immunoprophylaxis failure at amino acid level (AUC=0.677, 95% CI: 0.513 to 0.814, p=0.0247). (E) The ROC curve of the Shannon entropy in "a" determinant for predicting the risk of immunoprophylaxis failure at amino acid level (AUC=0.679, 95% CI: 0.516 to 0.816, p=0.0215). (F) The ROC curve of the quasispecies genetic distance in "a" determinant for predicting the risk of immunoprophylaxis failure at amino acid level (AUC=0.677, 95% CI: 0.513 to 0.814, p=0.0247). Blue curve represents the ROC curve of HBV DNA levels for predicting the risk of immunoprophylaxis failure (AUC=0.546, 95% CI: 0.383 to 0.702, p=0.639). Red dot indicated the point of the proposed optimum cut-off value. AUC: area under the curve; HBV, hepatitis B virus; NGS, next-generation sequencing; ROC: receiver operating characteristic.

Table 4 The adjusted ORs of immunoprophylaxis failure according to the cut-off values of these quasispecies characteristic in "a" determinant region using NGS assay

Characteristic (level)	Cut-off value	OR (95% CI)	β	SE	P value
Mutant frequency (nt)	0.0002	0.132 (0.029 to 0.594)	−2.022	0.766	0.008
Mutant frequency (aa)	0.0006	0.183 (0.042 to 0.810)	−1.696	0.758	0.025
Shannon entropy (nt)	0.0055	0.132 (0.029 to 0.594)	−2.022	0.766	0.008
Shannon entropy (aa)	0.0055	0.183 (0.042 to 0.810)	−1.696	0.758	0.025
Genetic distance (nt)	0.014	0.132 (0.029 to 0.594)	−2.022	0.766	0.008
Genetic distance (aa)	0.014	0.183 (0.042 to 0.810)	−1.696	0.758	0.025

β, regression coefficient; aa, amino acid; NGS, next-generation sequencing; nt, nucleotide.

level of "a" determinant region, respectively. Meanwhile, the AUC for immunoprophylaxis failure due to the baseline HBV DNA level was 0.546 (0.383–0.702, p=0.639), suggesting that baseline HBV DNA level could not predict the risk of immunoprophylaxis failure for their infants in the pregnant women with high HBV DNA load.

Moreover, the predictive values of these HBV quasispecies characteristics for immunoprophylaxis failure were also analysed by binary logistic regression assays. As shown in table 4, the cut-off values of these HBV quasispecies in predicting immunoprophylaxis failure were 0.0006 (mutant frequency), 0.0055 (Shannon entropy) and 0.014 (genetic distance) at amino acid level, and were 0.0002 (mutant frequency), 0.0055 (Shannon entropy) and 0.014 (genetic distance) at nucleotide level of "a" determinant region. According to the cut-off values of these HBV quasispecies characteristic, the corresponding adjusted ORs of immunoprophylaxis failure were (0.183, 95% CI: 0.042 to 0.810, p=0.025), (0.183, 95% CI: 0.042 to 0.810, p=0.025) and (0.183, 95% CI: 0.042 to 0.810, p=0.025) at amino acid level, and were (0.132, 95% CI: 0.029 to 0.594, p=0.008), (0.132, 95% CI: 0.029 to 0.594, p=0.008) and (0.132, 95% CI: 0.029 to 0.594, p=0.008) at nucleotide level of "a" determinant region, respectively (table 4).

DISCUSSION

Currently, the exact virological mechanism that leads to the immunoprophylaxis failure of infants born to mothers with high HBV DNA load remains a hot topic. Theoretically, the fidelity of HBV DNA polymerase is low during the reverse transcription of HBV pregenomic RNA; thus, the mutation rate during HBV replication is 10 times higher than that of other DNA viruses. Accordingly, HBV strains are also considered to be present as quasispecies in the HBV-infected individuals.[31 32] For HBV quasispecies, the characteristics of HBV quasispecies were generally analysed to demonstrate the evolution patterns of HBV under selective pressures, such as antiviral therapy, and the population characteristics of HBV among different hosts.[33 34] However, it has not yet reached a definitive conclusion on the role of quasispecies characteristics in the immunoprophylaxis failure of infants during HBV MTCT.

In this prospective nested case–control study, all the enrolled pregnant women did not receive antiviral treatment so that HBV strains could be better used to explain the virological mechanism

that leads to the immunoprophylaxis failure in infants. Since the short read of ultra-deep pyrosequencing was not suitable for analysing the full-length sequence of HBV genome,[35] the full-length HBV genome clone-based sequencing assay was considered the 'gold standard' for assessing HBV quasispecies characteristics. In this study, all of the full-length HBV genome clones derived from the same mother were clustered together, indicating that there was no contamination during the acquisition of clones. Furthermore, the sequences of full-length HBV genome in the mothers of two groups were evenly distributed in the phylogenetic tree, which indicated the good generalisability in the origin of HBV sequences and the similar HBV evolution pattern between the mothers of immunoprophylaxis failure infants and the mothers of immunoprophylaxis success infants.

In the mutant spectrum of quasispecies theory, the wild-type virus is no longer equated with a genome of a defined nucleotide sequence, but is defined as a steady-state mutant distribution dominated by a master sequence that displays the highest replication rate among the components of the mutant spectrum.[32 36] In this study, we selected HBV quasispecies as a unit to explore the relationship between the quasispecies characteristics and immunoprophylaxis failure. Interestingly, compared with the mothers of immunoprophylaxis failure infants, higher mutant frequency, Shannon entropy and mean genetic distance at amino acid level in "a" determinant region of HBV genome were found in the mothers of immunoprophylaxis success infants. Such phenomena were not due to the differences of a single site or some special sites in "a" determinant region between two groups (online supplementary figures S3–S6), even in the whole large HBsAg region (online supplementary figures S7–S10). Meanwhile, such phenomena were neither due to the differences of selection pressure of a single site nor the entire "a" determinant region between two groups (online supplementary table S8). Furthermore, NGS assay of "a" determinant region which could be sequenced within one reaction using Illumina Hiseq X confirmed the significant differences of the three quasispecies characteristics between two groups at amino acid level, as well as at nucleotide level. The observation at nucleotide level which was different from the clone sequencing assay (online supplementary figure S11, online supplementary table S9). Since a limited number of clones in each sample may be responsible for the absence of difference at nucleotides level, the huge reads of "a" determinant region in NGS assay might be closer to represent the real population than that in clone-based sequencing assay.

As we know, fitness represents a kind of ability to produce the stable infectious progeny in a given environment, which also includes viral replication capability, secretion, antigenicity and infectivity during the production of viral progeny.[37 38] It has been shown that a decreasing fitness of several viruses is always be resulted in bottleneck event.[39–41] MTCT of HBV is a population bottleneck event for HBV proliferation, which may result similarly in the decrease in HBV population fitness.[32 42] The reduction of virus population fitness can be compensated by the subsequent enlargement of virus population like multiple rounds of virus replication, whereas it may be needed to overcome the effect of Muller's ratchet.[32 43] Otherwise, high mutation rate will make virus population unable to maintain the dominant phenotypes of the quasispecies, and the virus population will enter into the error catastrophe and eventually extinct if the mutation rate is beyond threshold.[44 45] It has been proved that the enhanced mutagenesis and the increased quasispecies complexity can result in the reduction of virus infectivity or lead to virus extinction in vitro, which supported the above interpretations.[46 47] Accordingly, the new antiviral strategies are aimed to decrease the virus
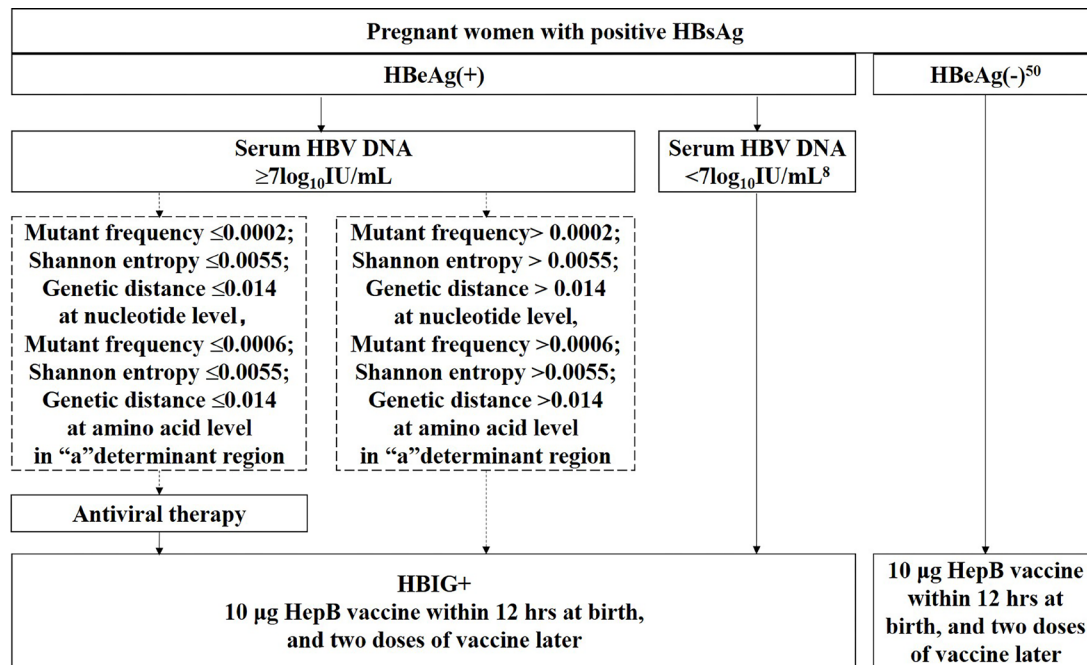
Figure 4 Management of pregnant women with positive HBsAg. HBIG, HBV immunoglobulin; HBV, hepatitis B virus; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen.

population fitness by increasing the lethal mutagenesis.[32 42] In this study, HBV quasispecies mutant frequency, Shannon entropy and mean genetic distance of "a" determinant region in the mothers of immunoprophylaxis success infants were higher than those in the mothers of immunoprophylaxis failure infants, indicating that the HBV quasispecies in the mothers of immunoprophylaxis success infants had the lower capacities of survival and infection through MTCT. It was suggested that HBV population with low fitness was difficult to overcome the Muller's ratchet and establish HBV infection in the infants with immunoprophylaxis, and the low fitness might work together with the deleterious effects of Muller's ratchet and error catastrophe for contributing to immunoprophylaxis success; thus, these HBV-infected pregnant women with low fitness of virus population might not need to receive antiviral therapy during the third trimester. Moreover, the ROC curve and the binary logistic regression analyses revealed that HBV quasispecies characteristics including mutant frequency, Shannon entropy and genetic distance in "a" determinant region were the potential predictors for the infants with high risk of immunoprophylaxis failure either at amino acid level or at nucleotide level. However, the predictive value of three quasispecies characteristics were higher at nucleotide level than that of amino acid level using NGS assay. We recognised one of the limitations in this study is the relatively small number of samples, which were derived from 1177 infants born to HBsAg-positive pregnant women without antiviral therapy. Therefore, large-scale multi-centre prospective studies are warranted to further confirm our findings in the future.

In conclusion, our study provides a new insight into the correlation of HBV quasispecies characteristics and immunoprophylaxis failure. Previous studies have shown that the mutations in "a" determinant region of HBV genome may induce conformational changes of HBsAg so that it may not be recognised by the neutralising anti-HBs antibodies and lead to immunoprophylaxis failure.[48 49] However, the immune escape mutations between the case group and control group, as well as the mother–child pairs were not found to be a significant factor of

immunoprophylaxis failure in this study (online supplementary tables S10–S15). Significantly, a more complex mutant spectrum in "a" determinant region was found in the mothers of immunoprophylaxis success infants, which indicated that HBV quasispecies in these mothers were more vulnerable to extinct during MTCT, and, in turn, lead to the eventual immunoprophylaxis success. Accordingly, the mutations in "a" determinant region of HBV genome might be a double-edged sword for the MTCT of the HBV. More importantly, the mutant frequency, Shannon entropy and genetic distance in "a" determinant region might be helpful to predict the infants with high risk of immunoprophylaxis failure in the pregnant women with high HBV DNA load. In conjunction with our previous study,[8 50] we proposed a flowchart (figure 4) for the management of pregnant women with positive HBsAg, which may be helpful to provide the precise medication to HBV-infected pregnant women with high HBV DNA load, and help the HBV-infected pregnant women who should not need antiviral therapy saving medical expenses and reducing the potential side effects of antiviral therapy.

Ethical approval The experimental protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki. This study was approved by the Ethics Committee of Peking University Health Science Center. Written informed consent was obtained from each subject.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iD
Yiwei Xiao http://orcid.org/0000-0003-0677-0746

REFERENCES

1 GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1736–88.

2 World Health Organization. Hepatitis B, 2018. Available: http://www.who.int/news-room/fact-sheets/detail/hepatitis-b

3 He T, Jia J. Chronic HBV: which pregnant women should be treated? *Liver Int* 2016;36(Suppl 1):105–8.

4 World Health Organization. Global health sector strategy on viral hepatitis 2016-2021, 2016. Available: http://apps.who.int/iris/bitstream/10665/246177/1/WHO-HIV-2016.06-eng.pdf

5 Dionne-Odom J, Tita ATN, Silverman NS, et al. #38: Hepatitis B in pregnancy screening, treatment, and prevention of vertical transmission. *Am J Obstet Gynecol* 2016;214:6–14.

6 Wiseman E, Fraser MA, Holden S, et al. Perinatal transmission of hepatitis B virus: an Australian experience. *Med J Aust* 2009;190:489–92.

7 Kubo A, Shlager L, Marks AR, et al. Prevention of vertical transmission of hepatitis B: an observational study. *Ann Intern Med* 2014;160:828–35.

8 Lu Y, Zhu F-C, Liu J-X, et al. The maternal viral threshold for antiviral prophylaxis of perinatal hepatitis B virus transmission in settings with limited resources: a large prospective cohort study in China. *Vaccine* 2017;35:6627–33.

9 Jaffe A, Brown RS. A review of antiviral use for the treatment of chronic hepatitis B virus infection in pregnant women. *Gastroenterol Hepatol* 2017;13:154–63.

10 Pan CQ, Duan Z, Dai E, et al. Tenofovir to prevent hepatitis B transmission in mothers with high viral load. *N Engl J Med* 2016;374:2324–34.

11 Jourdain G, Ngo-Giang-Huong N, Harrison L, et al. Tenofovir versus placebo to prevent perinatal transmission of hepatitis B. *N Engl J Med* 2018;378:911–23.

12 Hu Y, Xu C, Xu B, et al. Safety and efficacy of telbivudine in late pregnancy to prevent mother-to-child transmission of hepatitis B virus: a multicenter prospective cohort study. *J Viral Hepat* 2018;25:429–37.

13 Chen Z-X, Gu G-F, Bian Z-L, et al. Clinical course and perinatal transmission of chronic hepatitis B during pregnancy: a real-world prospective cohort study. *J Infect* 2017;75:146–54.

14 Zou H, Chen Y, Duan Z, et al. Virologic factors associated with failure to passive-active immunoprophylaxis in infants born to HBsAg-positive mothers. *J Viral Hepat* 2012;19:e18–25.

15 Carman WF, Zanetti AR, Karayiannis P, et al. Vaccine-Induced escape mutant of hepatitis B virus. *Lancet* 1990;336:325–9.

16 Coleman PF. Detecting hepatitis B surface antigen mutants. *Emerg Infect Dis* 2006;12:198–203.

17 Wang X, Deng W, Qian K, et al. Quasispecies characters of hepatitis B virus in immunoprophylaxis failure infants. *Eur J Clin Microbiol Infect Dis* 2018;37:1153–62.

18 Du Y, Chi X, Wang C, et al. Quantifying perinatal transmission of hepatitis B viral quasispecies by tag linkage deep sequencing. *Sci Rep* 2017;7:10168.

19 Wu Q, Xu C, Li J, et al. Evolution and mutations of hepatitis B virus quasispecies in genotype B and C during vertical transmission. *J Med Virol* 2016;88:1018–26.

20 Wei K-P, Zhu F-C, Liu J-X, et al. The efficacy of two different dosages of hepatitis B immunoglobulin combined with hepatitis B vaccine in preventing mother-to-child transmission of hepatitis B virus: a prospective cohort study. *Vaccine* 2018;36:256–63.

21 Centers for Disease Control and Prevention (CDC). Postvaccination serologic testing results for infants aged ≤24 months exposed to hepatitis B virus at birth: United States, 2008-2011. *MMWR Morb Mortal Wkly Rep* 2012;61:768–71.

22 Schillie SF, Murphy TV. Seroprotection after recombinant hepatitis B vaccination among newborn infants: a review. *Vaccine* 2013;31:2506–16.

23 Lewallen S, Courtright P. Epidemiology in practice: case-control studies. *Community Eye Health* 1998;11:57–8.

24 Grimes DA, Schulz KF. Compared to what? Finding controls for case-control studies. *Lancet* 2005;365:1429–33.

25 Günther S, Li BC, Miska S, et al. A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients. *J Virol* 1995;69:5437–44.

26 Kumar S, Stecher G, Li M, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–9.

27 Hao R, Xiang K, Peng Y, et al. Naturally occurring deletion/insertion mutations within HBV whole genome sequences in HBeAg-positive chronic hepatitis B patients are correlated with baseline serum HBsAg and HBeAg levels and might predict a shorter interval to HBeAg loss and seroconversion during antiviral treatment. *Infect Genet Evol* 2015;33:261–8.

28 Esteban D. *Quasispecies: concepts and implications for virology*, 2006.

29 Sun K-X, Li J, Zhu F-C, et al. A predictive value of quantitative HBsAg for serum HBV DNA level among HBeAg-positive pregnant women. *Vaccine* 2012;30:5335–40.

30 Chen J, Yan L, Zhu F-C, et al. Amino acid polymorphism in the reverse transcriptase region of hepatitis B virus and the relationship with nucleos(t)ide analogues treatment for preventing mother-to-infant transmission. *J Med Virol* 2014;86:1288–95.

31 Rodriguez-Frias F, Buti M, Tabernero D, et al. Quasispecies structure, cornerstone of hepatitis B virus infection: mass sequencing approach. *World J Gastroenterol* 2013;19:6995–7023.

32 Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev* 2012;76:159–216.

33 Liu F, Chen L, Yu D-M, et al. Evolutionary patterns of hepatitis B virus quasispecies under different selective pressures: correlation with antiviral efficacy. *Gut* 2011;60:1269–77.

34 Liu F, Yu D-M, Huang S-Y, et al. Clinical implications of evolutionary patterns of homologous, full-length hepatitis B virus quasispecies in different hosts after perinatal infection. *J Clin Microbiol* 2014;52:1556–65.

35 Gong L, Han Y, Chen L, et al. Comparison of next-generation sequencing and clone-based sequencing in analysis of hepatitis B virus reverse transcriptase quasispecies heterogeneity. *J Clin Microbiol* 2013;51:4087–94.

36 Schuster EDP. *Quasispecies: from theory to experimental systems*, 2016.

37 Domingo E, Holland JJ. Rna virus mutations and fitness for survival. *Annu Rev Microbiol* 1997;51:151–78.

38 Wargo AR, Kurath G. Viral fitness: definitions, measurement, and current insights. *Curr Opin Virol* 2012;2:538–45.

39 Joseph SB, Swanstrom R. Hiv/Aids. A fitness bottleneck in HIV-1 transmission. *Science* 2014;345:136–7.

40 de la Iglesia F, Elena SF. Fitness declines in tobacco etch virus upon serial bottleneck transfers. *J Virol* 2007;81:4941–7.

41 Chao L. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 1990;348:454–5.

42 Andino R, Domingo E. Viral quasispecies. *Virology* 2015;479-480:46–51.

43 Duarte EA, Clarke DK, Moya A, et al. Many-trillionfold amplification of single RNA virus particles fails to overcome the Muller's ratchet effect. *J Virol* 1993;67:3620–3.

44 Sardanyés J, Elena SF. Error threshold in RNA quasispecies models with complementation. *J Theor Biol* 2010;265:278–86.

45 Tejero H, Marín A, Montero F. The relationship between the error catastrophe, survival of the flattest, and natural selection. *BMC Evol Biol* 2011;11:2.

46 Grande-Pérez A, Sierra S, Castro MG, et al. Molecular indetermination in the transition to error catastrophe: systematic elimination of lymphocytic choriomeningitis virus through mutagenesis does not correlate linearly with large increases in mutant spectrum complexity. *Proc Natl Acad Sci U S A* 2002;99:12938–43.

47 Chung D-H, Sun Y, Parker WB, et al. Ribavirin reveals a lethal threshold of allowable mutation frequency for Hantaan virus. *J Virol* 2007;81:11722–9.

48 Salpini R, Colagrossi L, Bellocchi MC, et al. Hepatitis B surface antigen genetic elements critical for immune escape correlate with hepatitis B virus reactivation upon immunosuppression. *Hepatology* 2015;61:823–33.

49 Gencay M, Hübner K, Gohl P, et al. Ultra-deep sequencing reveals high prevalence and broad structural diversity of hepatitis B surface antigen mutations in a global population. *PLoS One* 2017;12:e0172101.

50 Lu Y, Liang X-F, Wang F-Z, et al. Hepatitis B vaccine alone may be enough for preventing hepatitis B virus transmission in neonates of HBsAg (+)/HBeAg (-) mothers. *Vaccine* 2017;35:40–5.