# The rs2296651 (S267F) variant on NTCP (SLC10A1) is inversely associated with chronic hepatitis B and progression to cirrhosis and hepatocellular carcinoma in patients with chronic…

**Article** *in* Gut · December 2015

DOI: 10.1136/gutjnl-2015-310686

CITATIONS
90

READS
554

**13 authors**, including:

Jessica Liu
Stanford Medicine
**65** PUBLICATIONS **1,983** CITATIONS
SEE PROFILE

Mei-Hsuan Lee
National Yang Ming Chiao Tung University
**181** PUBLICATIONS **7,721** CITATIONS
SEE PROFILE

Sheng-Nan Lu
Chang Gung Memorial Hospital
**568** PUBLICATIONS **23,434** CITATIONS
SEE PROFILE

San-Lin You
Academia Sinica
**259** PUBLICATIONS **19,739** CITATIONS
SEE PROFILE

ORIGINAL ARTICLE

# The rs2296651 (S267F) variant on NTCP (*SLC10A1*) is inversely associated with chronic hepatitis B and progression to cirrhosis and hepatocellular carcinoma in patients with chronic hepatitis B

Hui-Han Hu,[1] Jessica Liu,[1] Yu-Ling Lin,[1] Wun-Sheng Luo,[1] Yu-Ju Chu,[1] Chia-Lin Chang,[1] Chin-Lan Jen,[1] Mei-Hsuan Lee,[2] Sheng-Nan Lu,[3] Li-Yu Wang,[4] San-Lin You,[1] Hwai-I Yang,[1,2] Chien-Jen Chen,[1,5] for the REVEAL-HBV Study Group

▶ Additional material is published online only. To view this file please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2015-310686)

[1]Genomics Research Center, Academia Sinica, Taipei, Taiwan
[2]Institute of Clinical Medicine, National Yang-Ming University, Taipei, Taiwan
[3]Department of Gastroenterology, Chang-Gung Memorial Hospital, Kaohsiung, Taiwan
[4]Department of Medicine, Mackay Medical College, New Taipei City, Taiwan
[5]Graduate Institute of Epidemiology and Preventative Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

**Correspondence to**
Dr Hwai-I Yang, Genomics Research Center, Academia Sinica, 128 Academia Road Section 2, Nankang, Taipei 11529, Taiwan; hiyang@gate.sinica.edu.tw

H-HH and JL contributed equally.

Received 7 September 2015
Revised 7 November 2015
Accepted 9 November 2015

## ABSTRACT

**Objective** The sodium taurocholate co-transporting polypeptide (NTCP), encoded by *SLC10A1*, was recently identified as a receptor for HBV. We assessed the association of the p.Ser267Phe variant (rs2296651) with chronic hepatitis B (CHB) serostatus, cirrhosis and hepatocellular carcinoma (HCC) in patients with CHB.

**Design** The variant was genotyped in 3801 patients with CHB and 3801 matched hepatitis B surface antigen (HBsAg) seronegative individuals. ORs with 95% CIs for the variant's association with CHB, cirrhosis and HCC were estimated using logistic regression.

**Results** In patients with CHB, the S267F variant was observed in 515 (18.5%) controls, 40 (17.2%) cirrhosis only cases, 49 (13.2%) non-cirrhotic HCC cases, and 52 (12.7%) cirrhotic-HCC cases. After adjustment for known risk factors, S267F was significantly associated with decreased risk for cirrhosis (OR 0.65 (95% CI 0.49 to 0.86), p=0.002) and HCC (OR 0.55 (95% CI 0.42 to 0.72), p<0.001). This association persisted for non-cirrhotic and cirrhotic-HCC. Compared with patients with HBV DNA levels greater than $10^5$ copies/mL who carried the GG genotype, patients who had undetectable HBV DNA and the GA or AA genotypes had a 25-fold decreased risk of developing HCC (OR 0.04 (95% CI 0.02 to 0.11), p<0.001). The AA genotype was also associated with HBsAg seronegativity (OR 0.13 (95% CI 0.05 to 0.34), p<0.001).

**Conclusions** The *SLC10A1* (NTCP) S267F variant is independently associated with decreased risk of cirrhosis and HCC, and resistance to CHB infection. Together with serum HBV DNA levels, S267F may help to identify patients with CHB with very low risk of HCC.

## Significance of this study

### What is already known on this subject?

- The sodium taurocholate co-transporting polypeptide (NTCP), encoded by *SLC10A1*, was recently identified as a receptor for HBV.
- The p.Ser267Phe variant of *SLC10A1* is an Asian-specific variant which results in reduced HBV entry and infection.
- Previous association studies on NTCP S267F and HBV infection have yielded conflicting results and the role of this variant in cirrhosis and hepatocellular carcinoma (HCC) remains unclear.

### What are the new findings?

- The multivariate analyses adjusted for age, gender, alanine aminotransferase and HBV DNA levels show that the GA or AA genotypes of the S267F variant had significantly decreased risk for cirrhosis and HCC.
- The association of the S267F with HCC persisted in non-cirrhotic and cirrhotic HCC.
- Compared with patients with HBV DNA levels greater than $10^6$ copies/mL who carried the GG genotype, patients who had undetectable HBV DNA and the GA or AA genotypes had a 25-fold decreased risk of developing HCC.
- Contrasting patients with chronic hepatitis B (CHB) with non-HBV carriers, subjects who carried the S267F AA genotype were associated with HBsAg seronegativity.

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

- Together with serum HBV DNA levels, the S267F variant may help to identify a group of patients with CHB (5.5%) with very low risk of developing HCC in Asian communities of endemic HBV infection.

## INTRODUCTION

Chronic HBV infection is a serious public health problem worldwide due to its geographically widespread prevalence, particularly in the Asia-Pacific and sub-Saharan Africa regions, and its potential to cause liver cirrhosis, hepatocellular carcinoma (HCC) and death.[1 2] The natural history of chronic hepatitis B (CHB) is characterised by changes in seromarkers including alanine aminotransferase (ALT), hepatitis B e antigen (HBeAg), HBV DNA and hepatitis B surface antigen (HBsAg).[3] The seroclearance and transition rates of seromarkers have been studied,[4 5] and their impact on disease progression has also been evaluated recently.

**To cite:** Hu H-H, Liu J, Lin Y-L, *et al*. *Gut* Published Online First: [*please include* Day Month Year] doi:10.1136/gutjnl-2015-310686

Refined prediction models incorporating these seromarkers have been established to help predict the risk for cirrhosis and HCC among patients with CHB.[6 7]

HBV is an enveloped virus carrying a partially double-stranded DNA genome packaged in a nucleocapsid, which is shielded by a lipid bilayer containing small (S), medium (M) and large (L) surface glycoproteins.[8] The three proteins share a common C-terminal S domain, but the L and M proteins have different extensions of N-terminal domains. The N-terminal extensions of the L and M proteins are pre-S1/S2 and pre-S2, respectively.[9] The pre-S1 domain, especially amino acids 2–48, of the L protein plays a critical role in HBV entry and infection.[10 11] HBV exhibits prominent hepatotropism and only human and Tupaia hepatocytes are susceptible to HBV infection. A commonly accepted reason for this preference is specific receptor recognition of the viral pre-S1 domain.[12]

Recently, Yan *et al*[13] identified the sodium-dependent taurocholate co-transporting polypeptide (NTCP) as a cellular receptor for human HBV. NTCP is a transmembrane protein that is exclusively expressed on the basolateral membrane of differentiated hepatocytes, thus explaining the hepatotropism of HBV infection. The expression of NTCP is critical for HBV infection, and reduced expression of NTCP inhibits HBV infection of susceptible cells. Exogenous expression of human NTCP renders non-susceptible HCC cell lines susceptible to HBV infection.[13] Although monkey and mouse NTCPs are not functional receptors for HBV, replacing amino acids 157–165 of the monkey NTCP or 84–87 of the mouse NTCP with the corresponding parts of human NTCP, gives these NTCPs the ability to support HBV entry and infection.[13–15] These findings suggest that amino acids 157–165 and 84–87 are critical for human HBV entry and infection, also explaining the species-specificity of HBV infection.

NTCP is encoded by the *SLC10A1* gene, a member of the solute carrier family SLC10, located on chromosome 14. Several single nucleotide polymorphisms (SNPs) that alter the physiological function of NTCP have been identified and most of them show an ethnicity-dependent profile.[16 17] These functional SNPs would be expected to impact bile salt homoeostasis, liver function and HBV entry.

The p.Ser267Phe (S267F) variant (c.800G>A, rs2296651) on *SLC10A1* is an Asian-specific variant with a minor allele frequency (MAF) ranging from 3.1% to 9.2% among different Asian populations.[17] The S267F mutant of NTCP has normal cell surface expression but results in almost complete loss of bile salt uptake.[16–18]

Previous association studies on NTCP S267F and HBV infection have yielded conflicting results,[19 20] and the role of this variant in cirrhosis and HCC remains unclear. In addition, no studies have focused on identifying genetic variants in the key functional region between aa157 and 165 of NTCP among patients with CHB. We conducted a genetic association study in a large cohort of patients with CHB to assess the S267F variant in association with cirrhosis, HCC, and the seroclearance of HBeAg, HBV DNA and HBsAg. The *SLC10A1* gene was also partially sequenced to identify variants located in the aa157–165 region of NTCP. In addition, we compared this CHB cohort to another matched non-carriers cohort to assess the association between S267F variant and CHB serostatus.

## METHODS

### Study cohort

A total of 3801 CHB participants were used in this study, including 3292 participants from the REVEAL-HBV cohort, and 509 patients with HBV-related HCC obtained from the Taiwan Liver Cancer Network (TLCN). The REVEAL-HBV study is a community-based cohort of individuals aged 30–65 years at baseline who were HBsAg-positive, anti-HCV negative, and free of liver cirrhosis and HCC at recruitment during 1991–1992. Participants were followed up every 6–12 months with blood collection for future testing of HBV DNA, HBsAg, HBeAg and ALT, and did not receive antiviral treatment during the study period, as antiviral therapy was not reimbursed by the national health insurance for the majority of the study period and was only available under stringent criteria to higher-risk patients at the end of the study period. Detailed information on study procedures have been previously described.[21] The TLCN is comprised of five major medical centres across Taiwan, and aims to recruit patients with liver cancer from various socioeconomic and ethnic backgrounds. Participating centres follow standard protocol to collect tumour tissues and blood samples, as well as clinical, pathological and epidemiological information from patients with liver cancer. Samples were obtained from TLCN after application to the network and approval by the research committee. Patients from both cohorts were infected with HBV at birth, as in Taiwan, HBV is primarily transmitted perinatally. All participants from the REVEAL-HBV study or the TLCN provided written informed consent. In order to assess the association between S267F and CHB infection, another 3801 HBsAg-seronegative individuals were selected from the Taiwan Biobank, which were matched according to age and gender of patients with CHB with a 1:1 ratio. The Taiwan Biobank is a community-based cohort comprised of cancer-free Taiwanese participants aged between 30 years and 70 years with basic physical examination, questionnaire and genome-wide genotype data.

Patients were divided into four separate groups, as shown in figure 1. For the analysis of clinical outcomes, a total of 2789 individuals with no cirrhosis or HCC to be used as controls (Group I), 232 individuals who developed cirrhosis only (Group II), 370 individuals who developed non-cirrhotic HCC (Group III) and 410 individuals who developed cirrhotic-HCC (Group IV) were included. Analysis of serological outcomes was limited to patients in the REVEAL-HBV cohort. For the analysis of HBeAg seroclearance, the cohort was limited only to HBeAg seropositive individuals. For the analysis of HBV DNA and HBsAg seroclearance, the cohort was limited to HBeAg seronegative individuals with complete seroclearance data (figure 1).

### Ascertainment of outcomes

Patients in the REVEAL-HBV cohort were tested for HBeAg, HBsAg serostatus and HBV DNA levels every 6–12 months. Seroclearance was defined as the first instance in which an individual tested seronegative for HBeAg or HBsAg, or had undetectable serum HBV DNA levels (<100 copies/mL) and remained that way throughout the remainder of follow-up.

Cases of HCC and cirrhosis in the REVEAL cohort were detected through abdominal ultrasonography and α-fetoprotein testing, and through computerised linkages with the National Cancer Registry and National Death Certification databases. Identified cases were confirmed through chart reviews by gastroenterologists according to the following criteria: histopathological confirmation; positive lesions detected by at least two different imaging techniques (such as abdominal ultrasonography, angiogram or CT); or positive lesions detected by one imaging technique combined with a serum α-fetoprotein level greater than 400 ng/mL. Liver cirrhosis was determined with a

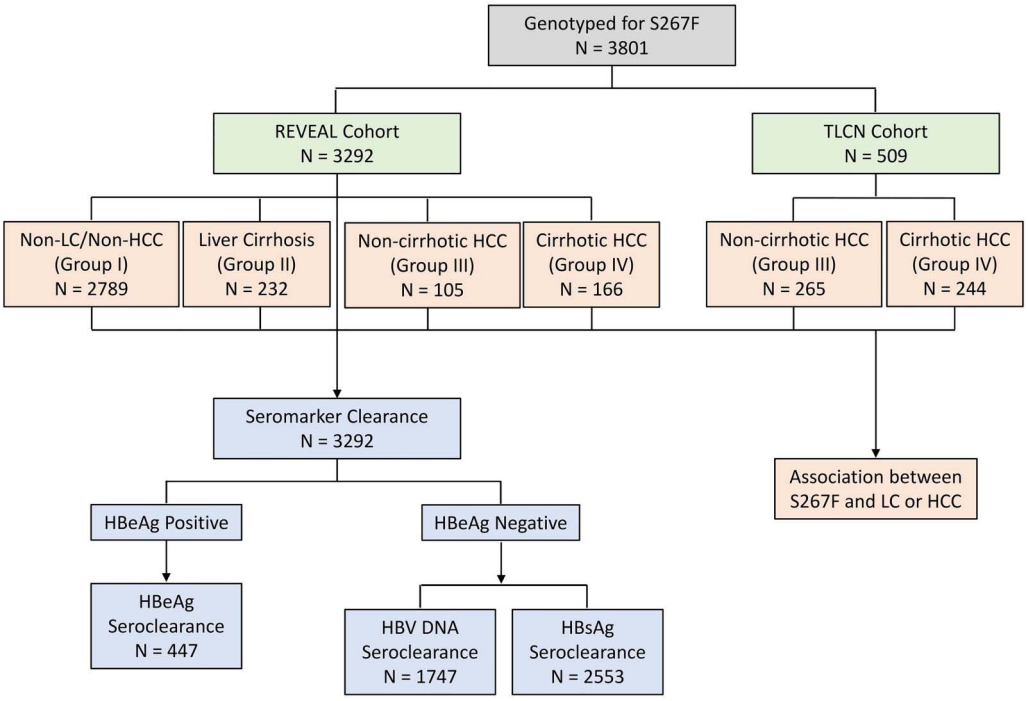


**Figure 1** Grouping strategy for the study population. Group I: Patients with chronic hepatitis B without LC and HCC; Group II: Patients with chronic hepatitis B who developed LC only; Group III: Patients with chronic hepatitis B who developed non-cirrhotic HCC; Group IV: Patients with chronic hepatitis B who developed cirrhotic HCC. HBeAg, hepatitis B e antigen; HCC, hepatocellular carcinoma; HBsAg, hepatitis B surface antigen; LC, liver cirrhosis; TLCN, Taiwan Liver Cancer Network.

quantitative scoring system based on liver parenchymal texture, intrahepatic blood vessel size and spleen size. In the TLCN cohort, HCC cases and cirrhosis diagnosis were ascertained through diagnosis and testing at participating hospitals.

### Genotyping and sequencing

Genomic DNA was extracted from buffy coat using the QIAamp DNA Blood Mini Kit (Qiagen, Valencia, California, USA) according to the manufacturer's protocol. The S267F variant (c.800G>A, rs2296651) was genotyped using the TaqMan genotyping assay (Applied Biosystems, Foster City, California, USA) (primer and probe sequences available on request). Exon 2 of the *SLC10A1* gene was first amplified by PCR using 20 ng of genomic DNA, deoxynucleotides (dNTPs) (0.2 mM each), the specific primer pair (0.4 µM each) (primer sequences available on request), 2 mM $MgCl_2$ and 0.1U of SuperTherm Taq (JMR Holdings, London, UK) in a final volume of 15 µL. The PCR was carried out at 95°C for 3 min, 35 cycles of 95°C for 30 s, 64°C for 30 s, and 72°C for 1 min, and finally, 72°C for 10 min. The PCR product was purified by the GenepHlow Gel/PCR kit (Geneaid Biotech). The PCR product was then sequenced by Sanger sequencing on the ABI 3730XL DNA analyser, using a BigDye Terminator V.3.1 Cycle Sequencing Kit (Applied Biosystems). Samples with variant calls from the first sequencing run were verified by a second run of the PCR using KAPA HIFI HotStart ReadyMix (Kapa Biosystems, Woburn, Massachusetts, USA). The PCR was carried out at 95°C for 5 min, 30 cycles of 98°C for 30 s, 62°C for 15 s, and 72°C for 30 s, and finally, 72°C for 7 min.

### In silico prediction tool

The effects of *SLC10A1* exon 2 variants on protein function were predicted by the following in silico prediction tools: SIFT (http://siftdna.org), SNPs3D (http://www.snps3d.org) and PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2). The mRNA splicing effect of intronic variants was predicted by Splice Site Prediction of Berkeley Drosophila Genome Project (BDGP) (http://www.fruitfly.org). Missense and intronic variants with functional impact predicted by at least one of four prediction tools were assigned as deleterious variants, while those without any functional impact were considered as non-deleterious variants. Nonsense mutations were directly classified as deleterious variants.

### Statistical analysis

Host and virological factors that were common to REVEAL and TLCN cohorts were included as predictors in outcome analyses. Differences in the distributions of host and virological characteristics between patient groups were compared with the $\chi^2$ test. Continuous variables were categorised for analysis. As detailed time-to-event data were not available in the TLCN cohort, all analyses of clinical outcomes (cirrhosis and HCC) were performed using unconditional logistic regression. Analysis of serological outcomes (ie, seroclearance of HBeAg, HBV DNA and HBsAg) was performed with Cox proportional hazards regression. Analysis of CHB infection was performed using univariate logistic regression, which compared HBsAg-seropositive patients with CHB from our cohort and a subset of HBsAg-seronegative individuals in the Taiwan Biobank with available S267F genotype data. All analyses were performed with SAS software V.9.3 (SAS Institute, Cary, North Carolina, USA), and statistical significance was determined using two-tailed tests ($p<0.05$).

## RESULTS

### Baseline characteristics of study cohort

As shown in table 1, a total number of 3801 participants were included in this study. The mean age (SD) was 46.4 years (10)

**Table 1** Baseline characteristics among four different groups of patients with chronic hepatitis B

| | Overall (n=3801) | Non-LC/non-HCC* (Group I) (n=2789) | LC case† (Group II) (n=232) | Non-LC HCC case‡ (Group III) (n=370) | LC-HCC case§ (Group IV) (n=410) |
|---|---|---|---|---|---|
| Age, mean years (SD) | 46.4 (10.0) | 45.1 (9.64) | 46.4 (9.89) | 49.8 (11.0) | 51.5 (9.44) |
| Gender | | | | | |
| Female | 1418 (37.3) | 1125 (40.3) | 56 (24.1) | 131 (35.4) | 106 (25.9) |
| Male | 2383 (62.7) | 1664 (59.7) | 176 (75.9) | 239 (64.6) | 304 (74.2) |
| Serum alanine aminotransferase level, U/L | | | | | |
| <45 | 3372 (88.7) | 2656 (95.2) | 200 (86.2) | 261 (70.5) | 255 (62.2) |
| ≥45 | 429 (11.3) | 133 (4.77) | 32 (13.8) | 109 (29.5) | 155 (37.8) |
| HBV DNA level, copies/mL | | | | | |
| <300 | 853 (22.4) | 769 (27.6) | 31 (13.4) | 31 (8.38) | 22 (5.37) |
| 300–9999 | 1127 (29.7) | 943 (33.8) | 49 (21.1) | 83 (22.4) | 52 (12.7) |
| 10 000–99 999 | 685 (18.0) | 498 (17.9) | 36 (15.5) | 79 (21.4) | 72 (17.6) |
| ≥100 000 | 1136 (29.9) | 579 (20.8) | 116 (50.0) | 177 (47.8) | 264 (64.4) |
| Serum HBsAg level, IU/mL¶ | | | | | |
| <100 | 939 (28.5) | 855 (30.7) | 31 (13.4) | 24 (22.9) | 29 (17.5) |
| 100–999 | 911 (27.7) | 800 (28.7) | 60 (25.9) | 19 (18.1) | 32 (19.3) |
| ≥1000 | 1442 (43.8) | 1134 (40.7) | 141 (60.8) | 62 (59.1) | 105 (63.3) |
| HBeAg serostatus¶ | | | | | |
| Positive | 491 (14.9) | 320 (11.5) | 55 (23.7) | 41 (39.1) | 75 (45.2) |
| Negative | 2801 (85.1) | 2469 (88.5) | 177 (76.3) | 64 (61.0) | 91 (54.8) |
| Viral genotype** | | | | | |
| C | 1034 (35.5) | 632 (31.2) | 80 (41.7) | 130 (41.1) | 192 (50.7) |
| B+BC | 1877 (64.5) | 1392 (68.8) | 112 (58.3) | 186 (58.9) | 187 (49.3) |
| Precore G1896A mutant†† | | | | | |
| Wild type | 478 (41.6) | 313 (37.1) | 64 (49.2) | 31 (50.8) | 70 (60.3) |
| Mutant | 672 (58.4) | 530 (62.9) | 66 (50.8) | 30 (49.2) | 46 (39.7) |
| rs2296651 (S267F) | | | | | |
| GG | 3145 (82.7) | 2274 (81.5) | 192 (82.8) | 321 (86.8) | 358 (87.3) |
| GA | 651 (17.1) | 510 (18.3) | 40 (17.2) | 49 (13.2) | 52 (12.7) |
| AA | 5 (0.13) | 5 (0.18) | 0 (0) | 0 (0) | 0 (0) |

*Patients with chronic hepatitis B without liver cirrhosis and HCC.
†Patients with chronic hepatitis B with liver cirrhosis but not yet developed HCC.
‡Patients with chronic hepatitis B who developed non-cirrhotic HCC.
§Patients with chronic hepatitis B who developed cirrhotic HCC.
¶Data were only available among the REVEAL-HBV cohort (n=3292).
**Data were available for individuals with detectable HBV DNA levels (≥300 copies/mL) in the REVEAL-HBV cohort.
††Data were only available among the REVEAL-HBV cohort with HBV DNA levels ≥10 000 copies/mL.
HBeAg, hepatitis B e antigen; HCC, hepatocellular carcinoma; HBsAg, hepatitis B surface antigen; LC, liver cirrhosis.

and the majority were male (62.7%), with serum ALT levels <45 U/L (88.7%), HBV DNA levels $<10^4$ copies/mL (52.1%) and serum HBsAg levels <1000 IU/mL (56.2%). Most of the patients were HBeAg seronegative (85.1%), infected with HBV genotype B or B+C (62.5%), carried the precore G1896A mutant (64.5%) and carried the GG genotype of the S267F variant (82.7%). The mean age of participants increased with disease progression (from Group I to Group IV), as did the proportion of participants with serum ALT levels ≥45 U/L, HBV DNA levels $\geq 10^4$ copies/mL, HBeAg seropositivity, HBV genotypes B and B+C, precore G1896A mutant and the S267F GG genotype.

The S267F variant was not statistically significantly associated with age, gender, serum ALT, HBV DNA, HBsAg levels and HBV genotype (see online supplementary table S1). Patients who carried an S267F variant (GA/AA) had a lower proportion of being infected by HBV with a stop codon mutation of HBeAg in the HBV precore region (G1896A), and a higher proportion of HBeAg seropositivity than those that carried a GG genotype.

### The association of S267F with liver cirrhosis

When patients who developed cirrhosis or cirrhotic-HCC during follow-up (Group II+Group IV) were compared with participants without cirrhosis and HCC (Group I), multivariate analyses showed that participants with the GA or AA genotype of the S267F variant had significantly decreased risk of liver cirrhosis (OR 0.65 (95% CI 0.49 to 0.86), p=0.002), after adjustment for age, gender, and serum ALT and HBV DNA levels (table 2).

### The association of S267F with HCC

To assess the association between S267F and HCC, participants who developed HCC during follow-up (Group III+Group IV) were compared with participants without cirrhosis and HCC (Group I) (table 2). Multivariate analyses adjusted for age, gender, serum ALT levels and HBV DNA levels showed that participants with the GA or AA genotype of the S267F variant had significantly decreased risk of HCC (OR 0.55 (95% CI 0.42 to 0.72), p<0.001). In addition, when we subdivided HCC into

**Table 2** Multivariate adjusted ORs of liver cirrhosis and hepatocellular carcinoma (HCC)‡

| | Liver cirrhosis | | | | HCC | | | |
|---|---|---|---|---|---|---|---|---|
| | Group II+Group IV vs Group I* | | | | Group III+Group IV vs Group I† | | | |
| | Control (n=2789) | Case (n=642) | OR (95% CI) | p Value | Control (n=2789) | Case (n=780) | OR (95% CI) | p Value |
| rs2296651 (S267F) | | | | | | | | |
| GG | 2274 (81.5) | 550 (85.7) | 1.00 | | 2274 (81.5) | 679 (87.1) | 1.00 | |
| GA/AA | 515 (18.5) | 92 (14.3) | 0.65 (0.49 to 0.86) | 0.002 | 515 (18.5) | 101 (13.0) | 0.54 (0.41 to 0.71) | <0.001 |
| Age, mean years (SD) | 45.1 (9.64) | 49.7 (9.90) | 1.07 (1.05 to 1.08) | <0.001 | 45.1 (9.64) | 50.7 (10.2) | 1.08 (1.07 to 1.09) | <0.001 |
| Gender | | | | | | | | |
| Female | 1125 (40.3) | 162 (25.2) | 1.00 | | 1125 (40.3) | 237 (30.4) | 1.00 | |
| Male | 1664 (59.7) | 480 (74.8) | 1.94 (1.56 to 2.42) | <0.001 | 1664 (59.7) | 543 (69.6) | 1.45 (1.18 to 1.77) | <0.001 |
| Serum alanine aminotransferase level, U/L | | | | | | | | |
| <45 | 2656 (95.2) | 455 (70.9) | 1.00 | | 2656 (95.2) | 516 (66.2) | 1.00 | |
| ≥45 | 133 (4.77) | 187 (29.1) | 4.99 (3.80 to 6.59) | <0.001 | 133 (4.77) | 264 (33.9) | 7.11 (5.47 to 9.23) | <0.001 |
| HBV DNA level, copies/mL | | | | | | | | |
| <300 | 769 (27.6) | 53 (8.26) | 1.00 | | 769 (27.6) | 53 (6.79) | 1.00 | |
| 300–9999 | 943 (33.8) | 101 (15.7) | 1.68 (1.18 to 2.40) | 0.004 | 943 (33.8) | 135 (17.3) | 2.35 (1.66 to 3.34) | <0.001 |
| 10 000–99 999 | 498 (17.9) | 108 (16.8) | 3.37 (2.35 to 4.84) | <0.001 | 498 (17.9) | 151 (19.4) | 4.74 (3.33 to 6.73) | <0.001 |
| ≥100 000 | 579 (20.8) | 380 (59.2) | 9.27 (6.68 to 12.86) | <0.001 | 579 (20.8) | 441 (56.5) | 10.9 (7.83 to 15.15) | <0.001 |

*Comparison of patients with liver cirrhosis to patients without liver cirrhosis and HCC.
†Comparison of patients who developed HCC to patients without liver cirrhosis and HCC.
‡HBsAg titre, HBeAg serostatus and HBV genotype were not included in this analysis due to the unavailability of these data in the Taiwan Liver Cancer Network (TLCN) cohort.

non-cirrhotic and cirrhotic-HCC (table 3), the association between the GA or AA genotype of the S267F variant with HCC remained significant, with an adjusted OR (95% CI) of 0.66 (0.47 to 0.92) for non-cirrhotic HCC and 0.48 (0.33 to 0.69) for cirrhotic HCC.

### HCC risk stratification according to viral factors and host S267F variant

To assess the interaction between S267F and HBV virological markers, associations with HCC were stratified by HBeAg serostatus, HBV DNA levels and HBsAg levels. The reduced risk of HCC associated with the S267F variant was not significantly different between HBeAg seropositive and seronegative patients (p=0.24), HBsAg levels <1000 IU/mL and ≥1000 IU/mL (p=0.32), or HBV DNA levels $<10^4$, $10^4$–$10^6$ and $\geq 10^6$ copies/mL (p=0.59) (data not shown). However, when compared with patients with HBV DNA levels $\geq 10^5$ copies/mL and the GG genotype, patients who had undetectable HBV DNA and the GA or AA genotypes of S267F had a 25-fold decreased risk for developing HCC (see online supplementary table S2; OR 0.04 (95% CI 0.02 to 0.10), p<0.001).

**Table 3** Multivariate adjusted ORs of non-cirrhotic hepatocellular carcinoma (HCC) and cirrhotic-HCC‡

| | Non-cirrhotic-HCC (Group III vs Group I)* | | | | Cirrhotic-HCC (Group IV vs Group I)† | | | |
|---|---|---|---|---|---|---|---|---|
| | Control (n=2789) | Case (n=370) | OR (95% CI) | p Value | Control (n=2789) | Case (n=410) | OR (95% CI) | p Value |
| rs2296651 (S267F) | | | | | | | | |
| GG | 2274 (81.5) | 321 (86.8) | 1.00 | | 2274 (81.5) | 358 (87.3) | 1.00 | |
| GA/AA | 515 (18.5) | 49 (13.2) | 0.65 (0.46 to 0.91) | 0.01 | 515 (18.5) | 52 (12.7) | 0.48 (0.33 to 0.69) | <0.001 |
| Age, mean years (SD) | 45.1 (9.64) | 49.8 (11.0) | 1.06 (1.05 to 1.08) | <0.001 | 45.1 (9.64) | 51.5 (9.44) | 1.10 (1.08 to 1.11) | <0.001 |
| Gender | | | | | | | | |
| Female | 1125 (40.3) | 131 (35.4) | 1.00 | | 1125 (40.3) | 106 (25.9) | 1.00 | |
| Male | 1664 (59.7) | 239 (65.6) | 1.07 (0.84 to 1.38) | 0.58 | 1664 (59.7) | 304 (74.2) | 2.03 (1.53 to 2.69) | <0.001 |
| Serum alanine aminotransferase level, U/L | | | | | | | | |
| <45 | 2656 (95.2) | 261 (70.5) | 1.00 | | 2656 (95.2) | 255 (62.2) | 1.00 | |
| ≥45 | 133 (4.77) | 109 (29.5) | 6.64 (4.84 to 9.10) | <0.001 | 133 (4.77) | 155 (37.8) | 7.68 (5.60 to 10.53) | <0.001 |
| HBV DNA level, copies/mL | | | | | | | | |
| <300 | 769 (27.6) | 31 (8.38) | 1.00 | | 769 (27.6) | 22 (5.37) | 1.00 | |
| 300–9999 | 943 (33.8) | 83 (22.4) | 2.45 (1.59 to 3.79) | <0.001 | 943 (33.8) | 52 (12.7) | 2.27 (1.34 to 3.83) | 0.002 |
| 10 000–99 999 | 498 (17.9) | 79 (21.4) | 4.18 (2.68 to 6.52) | <0.001 | 498 (17.9) | 72 (17.6) | 5.86 (3.51 to 9.79) | <0.001 |
| ≥100 000 | 579 (20.8) | 177 (47.8) | 7.42 (4.88 to 11.29) | <0.001 | 579 (20.8) | 264 (64.4) | 16.32 (10.11 to 26.36) | <0.001 |

*Multivariate analysis of the association between S267F and non-cirrhotic HCC was performed, comparing patients who developed non-cirrhotic HCC (Group III) to patients without liver cirrhosis and HCC (Group I).
†Multivariate analysis of the association between S267F and cirrhotic HCC was performed, comparing patients who developed cirrhotic HCC (Group IV) to patients without liver cirrhosis and HCC (Group I).
‡HBsAg titre, HBeAg serostatus and HBV genotype were not included in this analysis due to the unavailability of these data in the Taiwan Liver Cancer Network (TLCN) cohort.
HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen.

**Table 4** Association of S267F with chronic hepatitis B (CHB) infection

| | Patients with CHB* (n=3801) | HBsAg-seronegative controls† (n=3801) | OR (95% CI) | p Value |
|---|---|---|---|---|
| rs2296651 (S267) | | | | |
| GG | 3099 (81.53) | 3145 (82.74) | 1.00 | |
| GA | 665 (17.50) | 651 (17.13) | 0.96 (0.86 to 1.09) | 0.553 |
| AA | 37 (0.97) | 5 (0.13) | 0.13 (0.05 to 0.34) | <0.001 |

*Those are patients with CHB from our cohort for analyses in this work.
†The controls are a subset of HBsAg-seronegative individuals in the Taiwan Biobank with available S267F genotype data, which are matched according to gender and age of patients with CHB with a 1:1 ratio.
HBsAg, hepatitis B surface antigen.

### Association of S267F with CHB

When we compared our CHB cohort with the matched subset of HBsAg-seronegative individuals selected from the Taiwan Biobank, the AA genotype of the S267F variant was significantly and negatively associated with CHB infection (table 4; OR 0.13 (95% CI 0.05 to 0.34), p<0.001). The GA genotype of the variant also showed the same protective effect on CHB infection, although it was not statistically significant.

### Association of S267F with serological outcomes

The association of S267F with HBeAg seroclearance, HBV DNA undetectability and HBsAg seroclearance was also examined (see online supplementary table S3). However, after adjustment for age, gender, serum ALT levels, HBV DNA levels, HBsAg levels, HBV genotype and the HBV precore mutant, there was no significant association between the S267F variant and any serological outcomes.

### Rare variants

Sequencing of exon 2 of the *SLC10A1* gene identified no variants in the region between aa157 and 165 of NTCP. However, a total of 19 variants were identified in the same exon but exterior to the region between aa157 and 165 (see online supplementary table S4). The frequency of the majority of these variants was <1%, except for the p.Pro181Pro variant of which the frequency was 3.42%. Among 19 variants, 12 were novel variants that have never been described in the National Center for Biotechnology Information (NCBI) SNP database, the 1000 Genomes project and the Exome Variant Server. There were one nonsense, eight missense, seven silent and three intronic variants. Among eight missense variants, six were predicted to be deleterious, which may have functional protein effects, while one of the intronic variants was predicted to be a splice variant that could affect mRNA splicing. No association between these rare variants and cirrhosis and HCC was observed.

## DISCUSSION

In this study, we confirmed a high MAF of S267F (8.7%) in our community-based CHB cohort. We uncovered the protective association of the S267F variant of NTCP with cirrhosis and HCC (overall, cirrhotic and non-cirrhotic), and patients with GA or AA genotypes of S267F had lower risk of developing cirrhosis and HCC compared with those that carried GG genotype in each stratum of serum HBV DNA level. However, we failed to show the association between this genetic variant and the seroclearance of HBeAg, HBV DNA and HBsAg. To the best of our knowledge, this is the first association study demonstrating the protective association of the *SLC10A1* (NTCP) S267F variant with cirrhosis, cirrhotic-HCC and non-cirrhotic-HCC in a large cohort of patients with CHB.

Previous studies reported the MAF of S267F to be 7.5% in Chinese-Americans,[16] 3.1–5% in Koreans, 7.4% in Chinese individuals and 9.2% in Vietnamese individuals.[17] In our cohort, the MAF was similar to those from previous studies. However, the distribution of the minor allele was not in accordance with the Hardy-Weinberg equilibrium (p<0.001), due to less individuals being homozygous for S267F than the expected value. A study of limited number of subjects showed that in the Chinese Han population the S267F variant was associated with HBV infection (OR 2.91 (95% CI 1.10 to 7.74), p=0.03)[19]; while a larger cohort study of 1899 patients with HBV showed that S267F was associated with resistance to CHB (OR (95% CI) 0.36 (0.29 to 0.44) and 0.16 (0.06 to 0.40) for heterozygotes and homozygotes, respectively).[20] Our results showed that S267F AA homozygotes were strongly associated with HBsAg-seronegativity (OR 0.13 (95% CI 0.05 to 0.34), p<0.001), which support the findings of the later report.

The major physiological function of NTCP is the uptake of bile salts from portal blood into hepatocytes. The S267F variant is located in a highly conserved extracellular loop of NTCP which is important for bile salt interaction.[16 22] Although the S267F mutant shows normal cell surface expression, it results in almost non-existent bile salt uptake and reduced HBV infection.[16 23] Interestingly, cells cotransfected with wild type NTCP and the S267F mutant at a 1:1 ratio can be infected by HBV with an efficiency greater than 70%, suggesting that individuals carrying a single allele of the S267F variant are still susceptible to HBV infection.[23] This explains the presence of patients carrying the heterozygous S267F variant in our CHB cohort. Individuals who were homozygous for S267F were thought to lack a functional NTCP according to the in vitro study.[23] However, five patients with CHB in our cohort carried the homozygous S267F variant, and they had very low viral loads and HBsAg levels with none of them developing cirrhosis or HCC. Previous reports have found normal serum levels of bile salts in individuals homozygous for S267F, and bile salt uptake may be compensated by other transporters. [17] Therefore, the presence of patients with CHB with homozygous S267F may imply there exists minor alternative pathways for HBV entry. On the other hand, it is also possible that these individuals were infected by mutated forms of HBV. As the pre-S1 domain of HBV is highly polymorphic, variants or mutations of the HBsAg open reading frame could lead to conformational protein changes that affect binding with the S267F mutant of NTCP, resulting in HBV entry.

The mechanism underlying the association between the NTCP S267F variant and cirrhosis and HCC is still unknown. The NTCP is a bona fide HBV receptor, and binds and/or transports molecules including bile salts, steroid hormones, thyroid hormones, drug-conjugated bile salt and a variety of xenobiotics.[15] From our data, it seems unlikely to be through virological factors, as the variant was associated neither with most of baseline viral characteristics nor with the spontaneous clearance of viral seromarkers. Multivariate analyses clearly showed that even after adjustment for known virological risk factors, the association between S267F and cirrhosis and HCC remained significant, suggesting that the protective effect of S267F is independent of virological factors.

We hypothesised that the possible pathway underlying our observations may be through bile acids circulation. Bile acids are potentially cytotoxic compounds, and elevated concentrations

of bile acids within the liver trigger hepatocyte apoptosis by activating the death receptor pathway, thereby promoting liver injury and the development of liver cirrhosis and liver failure.[24] Several lines of evidence showed that bile acid concentrations that occur during cholestasis induce hepatocyte apoptosis, providing a cellular mechanism for bile acid-mediated liver injury.[24–26] The S267F mutant of NTCP may decrease the uptake of bile salts within the hepatocytes, reducing the possibility of intrahepatic cytotoxic bile salt accumulation. This may lower the likelihood of hepatic inflammation and oxidative stress-mediated oncogenesis in such patients with CHB, resulting in decreased risk of developing liver cirrhosis and HCC.

Yan *et al*[13] showed that aa157-165 of NTCP were critical for HBV binding and infection. We confirmed their results by not finding any genetic variants in the same region of NTCP in our CHB cohort. According to the 1000 Genomes database, only two variants exist in this region, but the allele frequency and population background of these variants are unknown. It has also been shown that alterations of residuals S162 and L163, which are conserved in human, mouse and monkey NTCPs, do not affect NTCP-mediated bile salt uptake but do reduce HBV infection.[23] These findings suggest that the region between aa157 and 165 of NTCP is important for HBV entry, and can be a useful target for new antiviral agent development.

Several molecules and substrates that target NTCP can inhibit HBV entry, and are expected to be potent candidates of anti-HBV drugs. Myrcludex-B mechanically competes with the natural L-protein for human NTCP binding,[15] and strongly inhibits HBV infection without severe impairment of bile salt uptake.[27] Ciclosporin A, an original immunosuppressive drug, its analogue SCYX1454139, and ezetimibe were reported to interact directly with NTCP and inhibit its bile salt uptake activity.[28–30] Other emerging compounds, including epigallocatechin-3-gallate, azelastine hydrochloride, ursodeoxycholate, cholate, progesterone and bromosulfophthalein, were reported to have anti-HBV entry activity.[31–33] Based on our findings, these agents may be useful for the inhibition of HBV infection, and hopefully for reducing risk of developing cirrhosis and HCC in patients with CHB. Furthermore, it would be interesting to study the response of patients carrying or not carrying the *SLC10A1* S267F variant treated with this new anti-HBV strategy.

There are several limitations to be considered. First, the ascertainment of cirrhosis based on ultrasound results only may not satisfactorily identify all patients with cirrhosis and hence, some individuals may be misclassified. Our finding on the association between S267F and cirrhosis should be conservative due to this non-differential misclassification. However, this issue may reduce the power for the analysis on non-cirrhotic HCC, as individuals with cirrhotic HCC may have been misclassified as non-cirrhotic HCC. Second, only exon 2 and the S267F SNP on the *SLC10A1* gene were sequenced and genotyped in our cohort. We did not conduct functional studies to assess the effect of predicted splice variant IV1-1G>A on mRNA splicing, or study the functional impact of the predicted deleterious rare variants identified on exon 2 of *SLC10A1*. As there are other SNPs and rare variants reported outside of this region, we may have underestimated the effect of *SLC10A1* (or NTCP) on CHB infection and HBV disease progression. According to studies on the effect of certain *SLC10A1* mutations on NTCP cell surface expression and substrate binding, it is reasonable to speculate that different conformations depending on post-translational modifications, membrane localisation and oligomerisation of NTCP[23] caused by other mutations or variants may have some impact on HBV entry.

In conclusion, in this large cohort of Taiwanese patients with CHB, we found that the *SLC10A1* (NTCP) S267F variant was significant and independently associated with cirrhosis, cirrhotic-HCC and non-cirrhotic-HCC, after adjustment of important risk factors including gender, age, and serum ALT, and HBV DNA levels. Together with serum HBV DNA levels, the S267F variant may help to identify a group of patients with CHB (5.5%) with very low risk of developing HCC in Asian communities of endemic HBV infection, and may also possibly be incorporated into cirrhosis and HCC prediction models for clinical use. Therapeutic agents targeting NTCP should be examined for their effect on the prevention of long-term outcomes including cirrhosis and HCC. Furthermore, we confirmed that this variant was associated with resistance to CHB infection.

**Contributors** The study was designed and supervised by H-IY, Y-LL, W-SL, Y-JC, C-LC and C-JC and C-LJ was in charge of sample preparation. H-HH, Y-LL and W-SL were responsible for experimental technical and material support. H-HH and JL conducted all statistical analyses, and H-HH, JL, H-IY and C-JC interpreted all results. H-HH drafted the manuscript. The manuscript was revised by JL, M-HL, H-IY and C-JC for important intellectual content. S-NL, L-YW and S-LY are members of the REVEAL-HBV study group and provided important material support. We thank the Taiwan Liver Cancer Network for providing the DNA samples and clinical data.

**Funding** The Department of Health, Taiwan; Bristol-Myers Squibb, USA; Roche Diagnostics, Switzerland; Academia Sinica, Taiwan; and Ministry of Science and Technology in Taiwan (101-2314-B-001-005-MY3).

**Competing interests** None declared.

**Ethics approval** Institutional Review Board of the College of Public Health, National Taiwan University, Taipei, Taiwan.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1 Custer B, Sullivan SD, Hazlet TK, *et al*. Global epidemiology of hepatitis B virus. *J Clin Gastroenterol* 2004;38(10 Suppl 3):S158–68.
2 Chen CJ, Wang LY, Yu MW. Epidemiology of hepatitis B virus infection in the Asia-Pacific region. *J Gastroenterol Hepatol* 2000;15(Suppl):E3–6.
3 Liaw YF, Chu CM. Hepatitis B virus infection. *Lancet* 2009;373:582–92.
4 Liu J, Yang HI, Lee MH, *et al*. Distinct seromarkers predict different milestones of chronic hepatitis B progression. *Hepatology* 2014;60:77–86.
5 Liu J, Lee MH, Batrla-Utermann R, *et al*. A predictive scoring system for the seroclearance of HBsAg in HBeAg-seronegative chronic hepatitis B patients with genotype B or C infection. *J Hepatol* 2013;58:853–60.
6 Yang HI, Lee MH, Liu J, *et al*. Risk calculators for hepatocellular carcinoma in patients affected with chronic hepatitis B in Asia. *World J Gastroenterol* 2014;20:6244–51.
7 Lee MH, Yang HI, Liu J, *et al*. Prediction models of long-term cirrhosis and hepatocellular carcinoma risk in chronic hepatitis B patients: risk scores integrating host and virus profiles. *Hepatology* 2013;58:546–54.
8 Liu J, Yang HI, Lee MH, *et al*. Spontaneous seroclearance of hepatitis B seromarkers and subsequent risk of hepatocellular carcinoma. *Gut* 2014;63:1648–57.
9 Baumert TF, Meredith L, Ni Y, *et al*. Entry of hepatitis B and C viruses—recent progress and future impact. *Curr Opin Virol* 2014;4:58–65.
10 Barrera A, Guerra B, Notvall L, *et al*. Mapping of the hepatitis B virus pre-S1 domain involved in receptor recognition. *J Virol* 2005;79:9786–98.
11 Glebe D, Urban S, Knoop EV, *et al*. Mapping of the hepatitis B virus attachment site by use of infection-inhibiting preS1 lipopeptides and tupaia hepatocytes. *Gastroenterology* 2005;129:234–45.
12 Schieck A, Schulze A, Gähler C, *et al*. Hepatitis B virus hepatotropism is mediated by specific receptor recognition in the liver and not restricted to susceptible hosts. *Hepatology* 2013;58:43–53.
13 Yan H, Zhong G, Xu G, *et al*. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *eLife* 2012;1:e00049.
14 Yan H, Peng B, He W, *et al*. Molecular determinants of hepatitis B and D virus entry restriction in mouse sodium taurocholate cotransporting polypeptide. *J Virol* 2013;87:7977–91.
15 Ni Y, Lempp FA, Mehrle S, *et al*. Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. *Gastroenterology* 2014;146:1070–83.

16 Ho RH, Leake BF, Roberts RL, *et al*. Ethnicity-dependent polymorphism in Na+-taurocholate cotransporting polypeptide (SLC10A1) reveals a domain critical for bile acid substrate recognition. *J Biol Chem* 2004;279:7213–22.
17 Pan W, Song IS, Shin HJ, *et al*. Genetic polymorphisms in Na+-taurocholate co-transporting polypeptide (NTCP) and ileal apical sodium-dependent bile acid transporter (ASBT) and ethnic comparisons of functional variants of NTCP among Asian populations. *Xenobiotica* 2011;41:501–10.
18 Watashi K, Urban S, Li W, *et al*. NTCP and beyond: opening the door to unveil hepatitis B virus entry. *Int J Mol Sci* 2014;15:2892–905.
19 Li N, Zhang P, Yang C, *et al*. Association of genetic variation of sodium taurocholate cotransporting polypeptide with chronic hepatitis B virus infection. *Genet Test Mol Biomarkers* 2014;18:425–9.
20 Peng L, Zhao Q, Li Q, *et al*. The p.Ser267Phe variant in SLC10A1 is associated with resistance to chronic hepatitis B. *Hepatology* 2015;61:1251–60.
21 Chen CJ, Yang HI. Natural history of chronic hepatitis B REVEALed. *J Gastroenterol Hepatol* 2011;26:628–38.
22 Mareninova O, Shin JM, Vagin O, *et al*. Topography of the membrane domain of the liver Na+-dependent bile acid transporter. *Biochemistry* 2005;44:13702–12.
23 Yan H, Peng B, Liu Y, *et al*. Viral entry of hepatitis B and D viruses and bile salts transportation share common molecular determinants on sodium taurocholate cotransporting polypeptide. *J Virol* 2014;88:3273–84.
24 Guicciardi ME, Gores GJ. Apoptosis: a mechanism of acute and chronic liver injury. *Gut* 2005;54:1024–33.
25 Miyoshi H, Rust C, Roberts PJ, *et al*. Hepatocyte apoptosis after bile duct ligation in the mouse involves Fas. *Gastroenterology* 1999;117:669–77.
26 Faubion WA, Guicciardi ME, Miyoshi H, *et al*. Toxic bile salts induce rodent hepatocyte apoptosis via direct activation of Fas. *J Clin Invest* 1999;103:137–45.
27 Petersen J, Dandri M, Mier W, *et al*. Prevention of hepatitis B virus infection in vivo by entry inhibitors derived from the large envelope protein. *Nat Biotechnol* 2008;26:335–41.
28 Watashi K, Sluder A, Daito T, *et al*. Cyclosporin A and its analogs inhibit hepatitis B virus entry into cultured hepatocytes through targeting a membrane transporter, sodium taurocholate cotransporting polypeptide (NTCP). *Hepatology* 2014;59:1726–37.
29 Nkongolo S, Ni Y, Lempp FA, *et al*. Cyclosporin A inhibits hepatitis B and hepatitis D virus entry by cyclophilin-independent interference with the NTCP receptor. *J Hepatol* 2014;60:723–31.
30 Lucifora J, Esser K, Protzer U. Ezetimibe blocks hepatitis B virus infection after virus uptake into hepatocytes. *Antiviral Res* 2013;97:195–7.
31 Huang HC, Tao MH, Hung TM, *et al*. (-)-Epigallocatechin-3-gallate inhibits entry of hepatitis B virus into hepatocytes. *Antiviral Res* 2014;111:100–11.
32 Kim RB, Leake B, Cvetkovic M, *et al*. Modulation by drugs of human hepatic sodium-dependent bile acid transporter (sodium taurocholate cotransporting polypeptide) activity. *J Pharmacol Exp Ther* 1999;291:1204–9.
33 Fu LL, Liu J, Chen Y, *et al*. In silico analysis and experimental validation of azelastine hydrochloride (N4) targeting sodium taurocholate co-transporting polypeptide (NTCP) in HBV therapy. *Cell Prolif* 2014;47:326–35.