particles were immunogenic and that insertion of a tag into the E2 protein could induce antibodies as a secondary immunogen.

4. Discussion

In this study we showed that infectious FLAG-tagged HCV particles with an N151K mutation that modulated HCV-glycosylation could be efficiently produced in cells and purified on a FLAG-agarose column. This purification procedure allowed analysis of the physical properties of the particles and the generation of anti-E2 antibodies.

The FLAG-tagged HCV particles were purified by simple anti-FLAG affinity chromatography in combination with ultrafiltration. However, the efficiency of purification was low, and the recovery of the HCV core protein in the final purified virus preparation was only approximately 5%. This low efficiency of purification of the HCV particles maybe due to a number of factors: (1) Interaction between the E2-FLAG protein and the anti-FLAG-agarose column may have been prevented by cellular host proteins with specific or non-specific affinity for the anti-FLAG-agarose, (2) A conformation change may have occurred in the E2 protein due to insertion of the FLAG sequence, which may have blocked FLAG interaction with anti-FLAG, (3) Free FLAG-E2 proteins may have bound more tightly to the FLAG-agarose than the FLAG-E2 protein on the viral surface. Nevertheless, sufficient purified FLAG-HCV particles were obtained with this purification procedure for further analysis.

In the density gradient analysis of the purified HCV particles, the peak of the HCV core protein coincided with the peaks of HCV RNA and viral infectivity. In previous reports of density gradient fractionation of non-FLAG-tagged viral particles, the peak of infectivity was reported to shift to a lighter density fraction than the peaks of HCV core protein or RNA [8,13,14]. This discrepancy between our, and previous, data may be explained by the fact that the properties of the purified FLAG-tagged HCV particles may differ from those of the JFH-1 based HCV particles reported previously. Regarding viral infectivity, it is known that cholesterol and sphingolipid association with HCV particles is important for virion maturation and infectivity [9]. The association of HCV particles with these lipids occurs in lipid rafts [9]. Since the E2-FLAG protein may have a decreased dependency on lipid rafts compared to the non-tagged E2 protein, this therefore resulted in a shift in the peak of infectivity. Alternatively, an association between HCV and very-low-density lipoprotein (VLDL) has an important role in HCV infectivity [15]. Tagging of HCV particles with FLAG may have somehow changed the association of HCV with VLDL and cause the observed shift in the infectivity peak. The mechanism of the shift of the infectivity peak of the FLAG-HCV particles needs to be examined in more detail in future studies.

HCV particles from human plasma samples have been previously examined by immunogold electron microscopy [16]. In the present study, we could clearly observe spherical particle structures of 40–60 nm in the purified samples. Furthermore, the FLAG-tagged HCV particles were aggregated by anti-FLAG antibody. This is the first report of the aggregation of HCV particles produced in an *in vitro* culture system. This method may therefore facilitate future examination of the detailed conformation of HCV particles and future elucidation of HCV particle structure by cryo-electron microscopy. However, it is regarded structural analysis as difficult to that aggregated HCV particles were inequable showing in this report because of this method also gathered defective viral particles which have the E2-FLAG protein.

Immunization of mice with purified FLAG-tagged HCV particles induced anti-E2 as well as anti-FLAG antibodies. These results indicated that the envelope proteins of the FLAG-tagged HCV particles were immunogenic and that insertion of a tag into the E2 protein could induce antibodies as a secondary immunogen. Thus, not only epitopes of viral origin, but also an epitope inserted into the virus, can induce an immune response. Although it is unclear how many amino acids can be inserted into the E2 HVR1, at least a triple FLAG-tag sequence (25 amino acids) is possible as shown in this study.

In conclusion, we have established a simple system for the purification of recombinant infectious FLAG-epitope-tagged HCV particles. The use of this system may contribute to studies aimed at a detailed analysis of HCV particle structure and towards HCV vaccine development.

Acknowledgments

This work was partially supported by a grant-in-aid for Scientific Research from the Japan Society for the Promotion of Science, from the Ministry of Health, Labor and Welfare of Japan, by the Research on Health Sciences Focusing on Drug Innovation from the Japan Health Sciences Foundation, and by the Japanese Society of Gastroenterology.

References

[1] K. Shimotohno, Hepatitis C virus as a causative agent of hepatocellular carcinoma, Intervirology 38 (1995) 162–169.
[2] Q.L. Choo, K.H. Richman, J.H. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, R. Medina-Selby, P.J. Barr, A.J. Weiner, D.W. Bradley, G. Kuo, M. Houghton, Genetic organization and diversity of the hepatitis C virus, Proc. Natl. Acad. Sci. USA 88 (1991) 2451–2455.
[3] N. Kato, M. Hijikata, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi, T. Sugimura, K. Shimotohno, Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis, Proc. Natl. Acad. Sci. USA 87 (1990) 9524–9528.
[4] Z. Stamataki, S. Coates, M.J. Evans, M. Wininger, K. Crawford, C. Dong, Y.L. Fong, D. Chien, S. Abrignani, P. Balfe, C.M. Rice, J.A. McKeating, M. Houghton, Hepatitis C virus envelope glycoprotein immunization of rodents elicits cross-reactive neutralizing antibodies, Vaccine 25 (2007) 7773–7784.
[5] T. Kato, T. Date, M. Miyamoto, A. Furusaka, K. Tokushige, M. Mizokami, T. Wakita, Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon, Gastroenterology 125 (2003) 1808–1817.
[6] T. Wakita, T. Pietschmann, T. Kato, T. Date, M. Miyamoto, Z. Zhao, K. Murthy, A. Habermann, H.G. Krausslich, M. Mizokami, R. Bartenschlager, T.J. Liang, Production of infectious hepatitis C virus in tissue culture from a cloned viral genome, Nat. Med. 11 (2005) 791–796.
[7] J. Zhong, P. Gastaminza, G. Cheng, S. Kapadia, T. Kato, D.R. Burton, S.F. Wieland, S.L. Uprichard, T. Wakita, F.V. Chisari, Robust hepatitis C virus infection in vitro, Proc. Natl. Acad. Sci. USA 102 (2005) 9294–9299.
[8] B.D. Lindenbach, M.J. Evans, A.J. Syder, B. Wolk, T.L. Tellinghuisen, C.C. Liu, T. Maruyama, R.O. Hynes, D.R. Burton, J.A. McKeating, C.M. Rice, Complete replication of hepatitis C virus in cell culture, Science 309 (2005) 623–626.
[9] H. Aizaki, K. Morikawa, M. Fukasawa, H. Hara, Y. Inoue, H. Tani, K. Saito, M. Nishijima, K. Hanada, Y. Matsuura, M.M. Lai, T. Miyamura, T. Wakita, T. Suzuki, Critical role of virion-associated cholesterol and sphingolipid in hepatitis C virus infection, J. Virol. 82 (2008) 5715–5724.
[10] M.J. van den Hoff, A.F. Moorman, W.H. Lamers, Electroporation in 'intracellular' buffer increases cell survival, Nucleic Acids Res. 20 (1992) 2902.
[11] T. Takeuchi, A. Katsume, T. Tanaka, A. Abe, K. Inoue, K. Tsukiyama-Kohara, R. Kawaguchi, S. Tanaka, M. Kohara, Real-time detection system for quantification of hepatitis C virus genome, Gastroenterology 116 (1999) 636–642.
[12] D. Delgrange, A. Pillez, S. Castelain, L. Cocquerel, Y. Rouille, J. Dubuisson, T. Wakita, G. Duverlie, C. Wychowski, Robust production of infectious viral particles in Huh-7 cells by introducing mutations in hepatitis C virus structural proteins, J. Gen. Virol. 88 (2007) 2495–2503.
[13] D. Akazawa, T. Date, K. Morikawa, A. Murayama, N. Omi, H. Takahashi, N. Nakamura, K. Ishii, T. Suzuki, M. Mizokami, H. Mochizuki, T. Wakita, Characterization of infectious hepatitis C virus from liver-derived cell lines, Biochem. Biophys. Res. Commun. 377 (2008) 747–751.
[14] Y. Miyanari, K. Atsuzawa, N. Usuda, K. Watashi, T. Hishiki, M. Zayas, R. Bartenschlager, T. Wakita, M. Hijikata, K. Shimotohno, The lipid droplet is an important organelle for hepatitis C virus production, Nat. Cell Biol. 9 (2007) 1089–1097.
[15] S.U. Nielsen, M.F. Bassendine, A.D. Burt, C. Martin, W. Pumeechockchai, G.L. Toms, Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients, J. Virol. 80 (2006) 2418–2428.
[16] M. Kaito, S. Watanabe, H. Tanaka, N. Fujita, M. Konishi, M. Iwasa, Y. Kobayashi, E.C. Gabazza, Y. Adachi, K. Tsukiyama-Kohara, M. Kohara, Morphological identification of hepatitis C virus E1 and E2 envelope glycoproteins on the virion surface using immunogold electron microscopy, Int. J. Mol. Med. 18 (2006) 673–678.

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

Evolution of hepatitis B genotype C viral quasi-species during hepatitis B e antigen seroconversion

Shuang Wu[1], Fumio Imazeki[1,*], Fuat Kurbanov[2], Kenichi Fukai[1], Makoto Arai[1], Tatsuo Kanda[1], Yutaka Yonemitsu[1], Yasuhito Tanaka[2], Masashi Mizokami[3], Osamu Yokosuka[1]

[1]*Department of Medicine and Clinical Oncology, Graduate School of Medicine, Chiba University, Chiba, Japan;* [2]*Department of Virology, Liver Unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan;* [3]*Research Center for Hepatitis and Immunology, International Medical Center of Japan Kounodai Hospital, Ichikawa, Japan*

Background & Aims: Although the evolution of viral quasi-species may be related to the pathological status of disease, little is known about this phenomenon in hepatitis B, particularly with respect to hepatitis B e antigen (HBeAg) seroconversion.

Methods: Nucleotide sequences of the hepatitis B virus (HBV) *X/precore/core* region was analyzed at five time-points in four groups of chronic hepatitis B patients, interferon-induced seroconverters (IS, $N = 9$), interferon non-responders (IN, $N = 9$), spontaneous seroconverters (SS, $N = 9$), and non-seroconverters (SN, $N = 9$) followed during 60 months on an average. Only patients with genotype C were studied.

Results: Analysis of 1800 nucleotide sequences showed that there was no statistical difference between the nucleotide genetic distances of seroconverters (IS and SS; 6.9×10^{-3} substitutions (st)/site and 6.7×10^{-3} st/site, respectively) and those of non-seroconverters (IN and SN; 5.3×10^{-3} st/site and 3.8×10^{-3} st/site, respectively) before seroconversion. Compared to non-seroconverters (IN and SN; 5.1×10^{-3} st/site and 5.9×10^{-3} st/site, respectively), the sequence diversity of seroconverters (IS and SS; 10.9×10^{-3} st/site and 9.9×10^{-3} st/site, respectively) was significantly higher after seroconversion ($p < 0.05$), and was higher in seroconverters after seroconversion than before seroconversion ($p < 0.05$), while this changed very little in non-seroconverters during the observation period. Phylogenetic trees showed greater complexity in secoconverters than non-seroconverters. Parsimony-based estimation of the direction of sequence change between descendants and ancestors before HBeAg seroconversion, revealed higher frequencies of transversional A to T substitution in seroconverters (0.06 *vs.* 0.02, $p = 0.0036$) that coincided with the dynamics of quasi-species possessing A1762T mutation.

Conclusions: The distinctly greater viral diversity in HBeAg seroconverters after seroconversion could be related to escape mutants resulting from stronger selection pressure.

Keywords: Chronic hepatitis B; Quasi-species; Hepatitis B e antigen seroconversion.

Received 15 December 2009; received in revised form 4 June 2010; accepted 7 June 2010; available online 25 August 2010

* Corresponding author. Address: Department of Medicine and Clinical Oncology, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-Ku, Chiba 260-8670, Japan. Tel.: +81 43 226 2083; fax: +81 43 226 2088.

E-mail address: imazekif@faculty.chiba-u.jp (F. Imazeki).

Abbreviations: SC, seroconversion; ALT, alanine aminotransferase; CHB, chronic hepatitis B; HBV, hepatitis B virus; IFN, interferon; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen; IS, interferon-induced HBeAg seroconverters; IN, IFN non-responders; SS, spontaneous seroconverters; SN, non-seroconverters.

Introduction

Hepatitis B virus (HBV) is a major human pathogen which can cause severe hepatic disease, including chronic hepatitis, cirrhosis (LC), and hepatocellular carcinoma (HCC). Quasi-species comprises a complex and dynamic distribution of non-identical but related genomes [1]. The evolution of viral quasi-species has been reported as important in the pathogenesis of RNA viruses such as hepatitis C virus [2–6] and human immunodeficiency virus [7–10], but little is known about HBV. HBV is a hepatotropic, non-cytopathic DNA virus replicated by an error-prone polymerase through an RNA intermediate. Because of this feature, the replication of HBV lacks fidelity. This results in a complex distributions of genomes with naturally-acquired mutations or mutations selected by either antiviral therapy or the immune response of the host. HBV quasi-species have not been subjected to detailed investigation, especially in the context of hepatitis B e antigen (HBeAg) seroconversion (SC), an immunologically mediated event. Whether there is a causal relationship between HBV seroconversion and HBV quasi-species remains unclear. HBV-related disease is known to be mediated both virologically and immunologically. Several studies have depicted the dynamic evolution of HBV quasi-species during lamivudine resistance or multiple drug resistance. This highlights the importance of HBV molecular evolution in revealing the mechanism of drug resistance [11,12]. HBV-specific cytotoxic T-cells play a significant role in the control of replication of HBV, which has been well documented in the literature [13–16].

Precore/core protein is the target of immunologically mediated HBeAg seroconversion. When the *precore/core* gene in HBV DNA is transcribed and translated, HBeAg is produced and secreted into the circulation [17,18]. But the synthesis and secretion of HBeAg are aborted by the emergence of a point mutation from G to A at nucleotide (nt)1896 (G1896A). Convincing lines of evidence have indicated a close association between HBeAg/anti-HBe seroconversion and the emergence of *precore* and *core* promoter mutations [19,20].

ELSEVIER

Journal of Hepatology **2011** vol. 54 | 19–25

Research Article

The purpose of this study was to elucidate the evolution of HBV quasi-species during HBeAg seroconversion. The results might help us to better understand the pathogenic mechanisms of HBV. We selected patients with well-characterized clinical phenotypes and compared their viral diversity based on the nucleotide sequences of the *X*/precore/core region. *Precore* and *core* promoter mutations were also investigated in detail before and after HBeAg seroconversion.

Materials and methods

Patients

Sera from 36 chronic hepatitis B patients with well-characterized clinical follow-up for >5 years were selected from a chronic hepatitis B database (77 seroconverters and 67 non-seroconverters) at Chiba University Hospital. Only patients with genotype C (subtype C2) were studied to ensure that differences found in viral evolution were not due to genotypic variation. Nine patients in each group were selected randomly if they fulfilled the following criteria and had sufficiently long follow-up. The index group comprised patients with documented HBeAg seroconversion (spontaneous seroconverters, SS), with serum at the following time-points relative to HBeAg seroconversion: time-point I (−25.2 ± 6.2/months), time-point II (−11.6 ± 2.7/months), time-point III (1 ± 2.3/months), time-point IV (12.5 ± 3.3/months), and time-point V (25 ± 3.6 months). Untreated control patients included those who were followed for a similar period of time and were persistently HBeAg positive (non-seroconverters, SN), and they were matched for average age of seroconversion and time-point intervals of the SS group. A second index group of patients with interferon (IFN)-induced HBeAg seroconversion (IFN seroconverters, IS), with serum at the following time-points relative to HBeAg seroconversion: time-point I (−24.3 ± 3.1/months), time-point II (−11.2 ± 1.9/months), time-point III (1 ± 1.2/months), time-point IV (12.7 ± 1.7/months), and time-point V (25.4 ± 2.2/months). Control patients were persistently HBeAg-positive despite IFN therapy (IFN non-responders, IN). Controls were matched for the average age of seroconversion, sex and time-point intervals of the IS group.

HBeAg seroconversion was defined as the loss of HBeAg and the development of anti-HBe. The serial serum samples in this study were taken at five time-points for each patient, as described above. This study was approved by the Ethics Committee of Chiba University Hospital.

Serological examination

HBsAg, HBeAg and anti-HBe were determined by enzyme-linked immunosorbent assay (ELISA; Abbott Laboratory, Chicago, IL). HBV genotype was determined from the patients' sera by ELISA (HBV genotype EIA; Tokushu-Meneki Laboratory, Tokyo, Japan), based on the method described by Usuda et al. [21]. Serum HBV DNA levels were monitored using the Roche Amplicor Monitor test (Roche Diagnostics, Tokyo, Japan), which has a lower detection limit of 2.6 $\log_{10}$ copies/ml, at each time-point.

Cloning and sequencing

Total DNA was extracted from 200 μl of each serum sample using QIAamp DNA blood mini kits (Qiagen, Hilden, Germany) according to the manufacturer's instructions and eluted in 200 μl distilled water. Because HBeAg seroconversion is associated with a decrease in HBV DNA levels, nested PCR was performed for all the samples. The primers for the first round of PCR were 5′-TCG CAT GGA GAC CAC CGT GA-3′ (sense, nt1604–1623) and 5′-ATA GCT TGC CTG AGT GC-3′ (antisense, nt 2076–2060). The primers for the second round of PCR were 5′-CAT AAG AGG ACT CTT GGA CT-3′ (sense, nt 1653–1672) and 5′-GGA AAG AAG TCA GAA GGC-3′ (antisense, nt 1974–1957).

Amplification was performed with 2 μl of DNA template (extracted DNA from serum samples for the first round PCR and the first round PCR products for the second round PCR) in 50 μl reaction under the following conditions: an initial 2 min of denaturation at 94 °C and 36 cycles of 94 °C denaturation for 1 min, annealing at either 54 °C or 52 °C for 1 min, in the first and second round respectively, and 72 °C extension for 1 min. The last cycle was followed by a final extension at 72 °C for 10 min. A 473-base pairs fragment (nt 1604–2076) containing the *X/precore/core* region was amplified.

PCR reactions were followed by cloning using TOPO® TA cloning kits (Invitrogen, Carlsbad, CA). All PCR products were purified with QIAquick PCR Purification Kit (Qiagen, Hilden, Germany), then cloned into the TOPO vector, and transformed into *Escherichia coli*. At least 15 clones per one cloning for samples from PCR reactions proceeded subsequent to the electrophoretic size separation on 1.2% agarose gel. Ten positive clones per cloning for samples from each PCR reaction were sequenced using BigDye® Terminator and a 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA). The cloning PCR and sequencing primers were M13-forward, 5′-GTA AAA CGA CGG CCA GT-3′, and M13-reverse, 5′-GGA AAC AGC TAT GAC CAT G-3′.

Sequence analysis

The DNAPARS program from PHYLIP v3.65 package, implemented in Simmonic Sequence Editor version 1.5 [22], was used for sequence analysis. To evaluate quasi-species-based evolution of HBV strains in chronic patients, sequences of clones ($N = 10$) isolated at each time-point ($N = 5$) from individual patients ($N = 36$) were subjected to alignment and used to generate one parsimonious ancestral sequence. Maximum nucleotide composition distances were evaluated pair-wise between the ancestral sequence and the sequences of each of the 10 clones with a mean value estimated for each patient at a given time-point (*MEGA* version 4 [23]). All patients were categorized into four groups with respect to seroconversion status and the mean distance value for each group was calculated for each time-point.

The differences in genetic distance among clinical groups and time-points, and diversity at each time-point, were analyzed using ANOVA (analysis of variance). Student's *t*-test was also performed to determine the average of genetic diversities in non-seroconverters. All graphical data are presented as means ± standard deviation (SD). Results were considered statistically significant at $p < 0.05$. The statistical analysis was performed with SPSS (2004; SPSS Inc., Tokyo, Japan).

Construction of phylogenetic trees

To examine the evolution of the viral sequence and whether this evolution was elicited by quasi-species or mutagenesis, phylogenetic trees were constructed using the Neighbor-Joining (NJ) model with the Simmonic Sequence Editor version 1.5, based on the genomic sequences of HBV. Moreover, to investigate viral genetic features possibly associated with seroconversion, sequences isolated at time-points 1 and 2 were further analyzed phylogenetically. Neighbor-Joining trees were constructed at time-points 1 and 2 (Fig. S1 and S2, respectively) using all groups of sequences.

Results

Baseline clinical characteristics of the patients and sequential levels of serum ALT and HBV DNA

The clinical and laboratory characteristics of all patients are listed in Table 1. The levels of alanine aminotransferase (ALT) and HBV DNA over time are illustrated in Fig. 1A and B, respectively. Serum ALT levels, a marker of hepatocyte damage, normalized after seroconversion and, for all groups except the interferon non-responders, were <40 IU/L at the end-point of observation. HBV DNA loads decreased markedly in seroconverters (<3 $\log_{10}$ copies/ml, $p < 0.0001$) but changed very little in non-seroconverters. It is noteworthy that, at the second year after seroconversion, serum HBV DNA loads increased in interferon-induced seroconverters compared to spontaneous seroconverters, without statistical significance ($p^H = 0.1087$) (Fig. 1B).

Viral nucleotide sequence diversity

Viral sequence diversity, phylogenetic trees, and mutation pattern based on 1800 HBV nucleotide sequences from clones of the *X/precore/core* region, were analyzed among selected patients.

Table 1. Baseline clinical features of patients.

	IFN Seroconverters (IS)	IFN Non-seroconverters (IN)	Spontaneous Seroconverters (SS)	Spontaneous Non-seroconverters (SN)
Age (y)	40 ± 8	40 ± 11	29 ± 10	34 ± 6
Male : Female	6:3	8:1	5:4	7:2
HBV DNA (log_{10} copies/ml)	6.8 ± 0.9	6.8 ± 1.0	6.8 ± 1.2	7.1 ± 0.8
ALT (IU/L)	88.3 ± 48.6	94.3 ± 144.4	89.8 ± 71.4	67.6 ± 48.7

Note: The IFN-induced group (seroconverters and non-responders) was older than the spontaneous group (seroconverters and non-responders). Males were the majority in all groups. Baseline serum HBV DNA and ALT levels are similar among the four groups. Data are shown as mean ± SD.

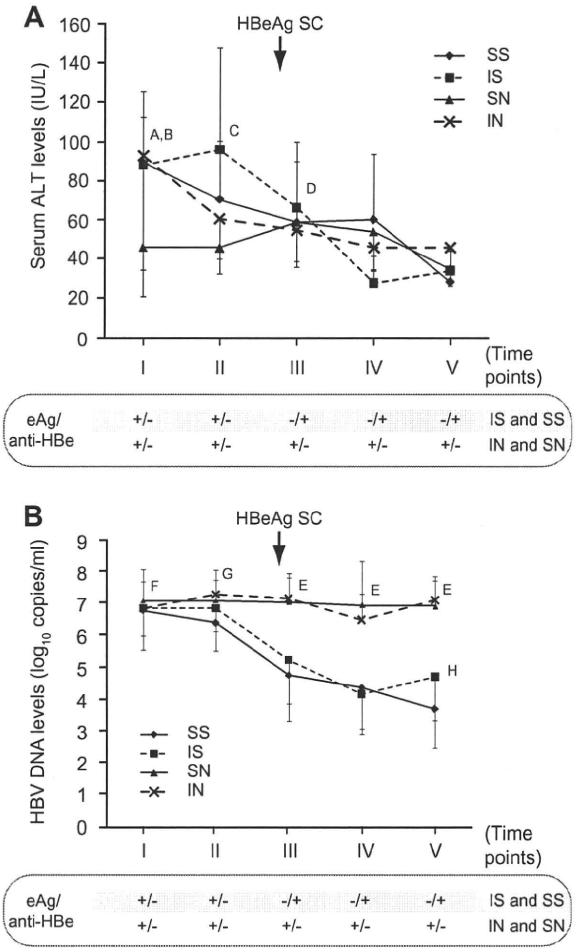

Fig. 1. Serum ALT and DNA levels in the four groups. The group of spontaneous seroconverters (SS) is a solid line diamond, IFN-induced seroconverters (IS) is a broken line square, IFN non-responders (IN) is a broken line asterisk, and non-seroconverters controls (SN) is a solid line triangle. (A) $p^A = 0.0234$ comparing time-point I with time-point IV, $p^B = 0.0028$ comparing time-point I with time-point V, $p^C = 0.007$ comparing time-point II with time-point V, $p^D = 0.0068$ comparing time-point III with time-point V. (B) $p^E < 0.0001$ comparing seroconverters with non-seroconverters, $p^F < 0.0001$ comparing time-point I with III, IV, V, $p^G < 0.0001$ comparing time-point II with the other time-points, $p^H = 0.1087$ at time-point V in seroconverters.

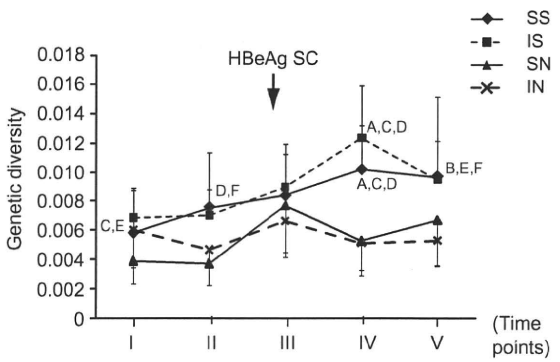

Fig. 2. Viral genetic diversity in the four groups. The group of spontaneous seroconverters (SS) is a solid line diamond, IFN-induced seroconverters (IS) is a broken line square, IFN non-responders (IN) is a broken line asterisk and non-seroconverters controls (SN) is a solid line triangle. $p^A < 0.0001$ comparing seroconverters with non-seroconverters at time-point IV, $p^B = 0.0301$ comparing seroconverters with non-seroconverters at time-point V, $p^C = 0.0013$ and $p^D = 0.0025$ comparing I and II with time-point IV in seroconverters. $p^E = 0.0121$ and $p^F = 0.021$ comparing time-points I and II with V in seroconverters.

Striking differences in nucleotide sequence diversity were revealed between seroconverters and non-seroconverters before and after seroconversion (Fig. 2). The nucleotide sequence diversity of seroconverters was similar to that of non-seroconverters before seroconversion. Analysis of genetic distance showed that the viral sequence diversity of seroconverters was significantly greater than that of non-seroconverters after seroconversion (Fig. 2, $p^A < 0.0001$ at time-point IV, $p^B = 0.0301$ at time-point V) and was greater in seroconverters after seroconversion than before (Fig. 2, $p^C = 0.0013$ and $p^D = 0.0025$), while almost no changes were observed in non-seroconverters during the observation period.

It is noteworthy that, in interferon-induced seroconverters at the last time-point of observation, the nucleotide sequence diversity was less, although this increased clearly at the first year after seroconversion. This tendency of reversed change at the last two time-points was also seen in HBV DNA loads (Fig. 1B), namely, increase or decrease of the genetic diversity accompanied by decrease or increase of the viral load in interferon-induced seroconverters. On the other hand, the nucleotide sequence diversity increased continuously in spontaneous seroconverters, accompanied by a concurrent decrease of viral loads (Fig. 1B) during the follow-up period. Amino acid sequence diversity had an almost

Research Article

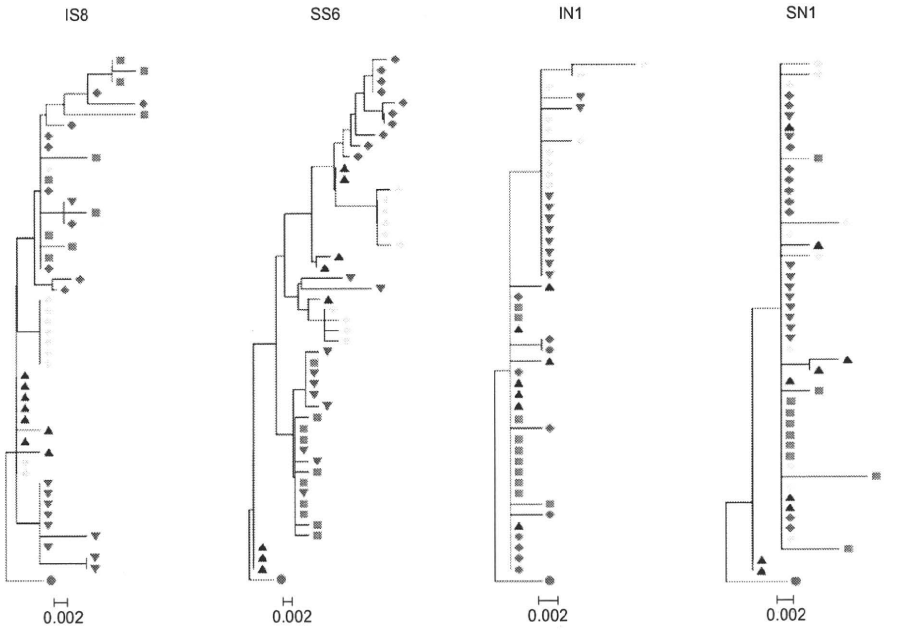

Fig. 3. Representative Neighbor-Joining phylogenetic trees of HBV sequences for each clinical group showing complex trees in seroconverters. HBV *X/precore/core* sequences from time-points I (purple filled triangle), II (blue filled inverted triangle), III (green filled square), IV (red filled diamond) and V (sky blue filled diamond) serum samples are analyzed phylogenetically and their positions are displayed on the trees. A sequence retrieved from the time-point I (red dot) of each group as outgroup in the trees, respectively. Scale bar represents 0.002% genetic variation. Seroconversion patients (IS, SS) show relatively complex branching patterns, forming clusters over time. With the pressure of seroconversion, the genetic diversity increased. In contrast, patients without seroconversion (IN, SN) were simply branching patterns and the genetic diversity in these patients changed very little over time.

identical pattern to that of DNA nucleotide sequence diversity (data not shown).

Construction of phylogenetic trees

Phylogenetic trees were complex for secoconverters and comparatively simple for non-seroconverters. In seroconverters (IS and SS), the arrangement and branch lengths of the trees were consistently more complex and longer than those for non-seroconverters. The genetic diversity was great after seroconversion in seroconverters (IS and SS) and less in non-seroconverters (IN and SN) (Fig. 3).

To investigate viral genetic features possibly associated with seroconversion, sequences isolated at time-points 1 and 2 (before seroconversion) were further analyzed phylogenetically. Trees were reconstructed using Neighbor-Joining, ML (data not shown), and PAML methods (data not shown). In general, no clusters were seen to be supported by robust bootstrap values for any group or particular patient quasi-species. This indicates that the region of the HBV genome studied does not contain patterns of variability sufficient for robust phylogenetic relation reconstruction. However, variability of branch lengths in the tree indicated that seroconversion patient groups exhibit greater diversity of the quasi-species compared to patients without seroconversion. This is in agreement with the genetic distance plot (Fig. 2), showing greater deviation from the mean values observed in patients with seroconversion. The IN group exhibited least deviation on the distance plot (Fig. 2) and shortest branch lengths on the trees (Fig. 3).

Interclonal differences of the quasi-species

To investigate whether a particular mutation pattern of evolution of the quasi-species is associated with seroconversion, we further analyzed the sequence changes in all patients at time-points 1 and 2, corresponding to the time before seroconversion. Parsimony-based ancestral sequences were generated using the Simmonic Sequence Editor. Aligned sequences of time-points 1 and 2 from a single patient were used as the input. Frequencies of changes in 12 types of mutations, including 4 transitions (CT, TC, AG, and GA) and 8 transversions (AT, TA, AC, CA, CG, GC, GT, and TG) were evaluated between generated descendants and ancestral sequences for each clone of the patient. Statistical *t*-test comparison of mean values of nucleotide changes between seroconversion and non-seroconversion groups is summarized in Table 2 and Supplementary Table 1.

Analysis of sequence changes indicated a higher frequency of transversional A to T in spontaneous seroconverters (SS *vs.* SN = 0.06 *vs.* 0.02, $p = 0.04$) and IFN-induced seroconverters (IS *vs.* IN = 0.05 *vs.* 0.01, $p = 0.05$) and A to C changes in IFN-induced seroconverters (IS *vs.* IN = 0.025 *vs.* 0.006, $p = 0.04$) before seroconversion. Comparison of seroconversion groups (SS and IS) indicated a higher frequency of transversional A to T mutation pattern ($p = 0.003$, Table 2) and the trend of G to A mutation is higher in seroconversion groups (SS and IS) (Table 2). Subsequently, alignments of the clones were generated. Visual inspection of the alignments indicated variation in the ratio of A1762T mutation in clones isolated from each patient at time-points 1 and 2 (Fig. 4). In contrast to non-seroconverters, seroconverters

Table 2. *t*-test comparison of mean values of nucleotide changes between seroconversion and non-seroconversion groups.

	Seroconversion (n = 18)	Non-seroconversion (n = 18)	*p*
CT	0.117033	0.103750	0.637023
TC	0.156706	0.201328	0.155252
AG	0.125483	0.148372	0.498916
GA	0.196722	0.124511	0.073433
AT	0.061194	0.022128	0.003665
TA	0.049372	0.045417	0.778612
AC	0.027944	0.012550	0.145158
CA	0.017128	0.011094	0.523868
CG	0.009439	0.007744	0.835337
GC	0.018167	0.014894	0.748267
GT	0.009839	0.019217	0.272185
TG	0.041783	0.035528	0.731324

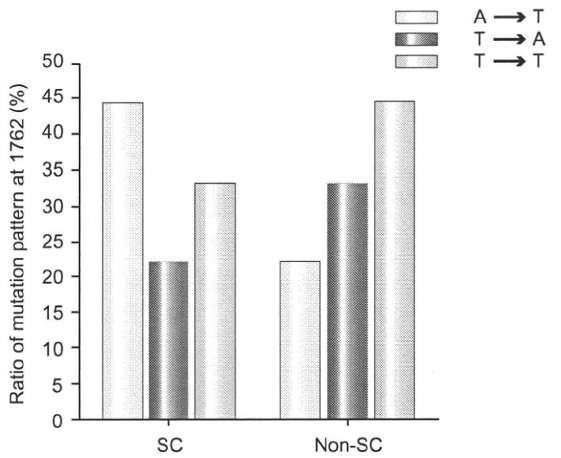

Fig. 4. The evolution of the *core* promoter mutation (A1762T) between seroconversion and control groups from time-point I to II. SC indicates seroconversion and non-SC, non-seroconversion. Alignment of the clones was carried out and the frequency of A1762T mutation in clones isolated from each patient at time-points 1 and 2 was determined. Subsequently, the evolutionary ratio of mutation from time-point I to II was calculated.

showed a higher frequency of A to T mutation pattern in the core promoter region from time-point I to II.

Core promoter (A1762T/G1764A) and precore (G1896A) mutations

Given that the *core* promoter/*precore* mutations influenced virus replication and HBeAg seroconversion, we analyzed the sequential change of *core* promoter (A1762T/G1764A)/*precore* (G1896A) mutations over time (Table 3). After seroconversion, patients with more than 50% *precore* mutant clone had higher HBV DNA loads than those with less than 50% of *precore* mutant clone (*precore* wild type) virus at time-point V [5.4 ± 1.3 (n = 5) *vs.* 3.8 ± 1.1 (n = 13), p = 0.0185] and 8 patients with a HBV DNA load less than 4.0 $\log_{10}$ copies/ml had all *precore* wild-type virus at time-point V (Table 3). Clinical progress of these patients was investigated over 10 years as median (range 1–20 years) after HBeAg seroconversion. HCC developed in 3 of 5 patients with *precore* mutant virus, compared to 1 of 13 patients with *precore* wild-type virus at time-point V (p = 0.017). On the other hand, 3 patients with ASC had all *precore* wild-type virus at time-point V (Table 3).

Discussion

In this study, analysis of 1800 nucleotide sequences from 36 HBV carriers showed that the viral diversity of seroconverters (IS and SS) after seroconversion was significantly greater than that of non-seroconverters (IN and SN) (Fig. 2, p <0.05) and was higher after seroconversion than before, in the seroconverters (Fig. 2, p <0.05). Phylogenetic analysis also generated complex trees for seroconverters and relatively simple trees for non-seroconverters. Analysis on interclonal differences in the quasi-specie showed a higher frequency of transversional A to T mutation pattern in seroconverters that coincided with the A1762T core promoter mutation. These findings suggested that HBeAg seroconversion involves dynamic shifts of the serum HBV quasi-species.

Osiowy et al. [24] examined viral quasi-species in eight HBeAg-negative patients at two time-points 25 years apart and obtained the evolutionary rate. Their results suggested that HBV diversity may be generated more rapidly than those estimated previously [25–29]. The higher evolutionary rate may be related to the seroconversion event driving quasi-specie complexity and diversification [24]. Our phylogenetic study showed that viral quasi-species populations appear to be replaced by new populations arising from a different clade after seroconversion.

Increased immune responses are accompanied by the reduction of viral loads and stronger immune pressure induces the selection of escape mutations, which leads to greater viral diversity [30]. According to this scenario, in our study, non-seroconverters have a high viral load and low quasi-species diversity and they obviously have a weak immune response.

Lim et al. [31] reported that viral genetic diversity in genotype B CHB patients was 2.4-fold greater in HBeAg seroconverters (spontaneous or IFN-induced) than in non-seroconverters before seroconversion. In this study of genotype C CHB patients, the nucleotide genetic distance was 1.49-fold greater in seroconverters (IS and SS) than in non-seroconverters before seroconversion but there was no statistical difference. This discrepancy might be due to the smaller region for analysis of genetic distance in our study than that of Lim et al. Another interpretation is that the host's immune response to the selection of mutant virus might differ between genotype B and genotype C. The natural course of CHB and the response to treatment could be affected by HBV genotype and there are some lines of evidence that indicate that the prevalence rates of *precore* and *core* promoter mutations vary among patients infected with HBV strains of different genotypes [32–34].

T-test comparison of mean values of nucleotide changes (Table 2) and linear logistic regression univariate analysis of mutations associated with seroconversion between seroconverters and non-seroconverters (data not shown) indicated a variation in the AT mutation pattern in the former (p = 0.003 and p = 0.006, respectively). This coincided with differences in the

Table 3. *Core* promoter and *precore* mutations in seroconverters (IS and SS).

Patients	CP (ntA1762T/G1764A) (percent)			PC (ntG1896A) (percent)			DNA Loads (log_{10} copies/ml)			Histological diagnosis
	I	III	V	I	III	V	I	III	V	
IS1	100	100	100	0	0	70	5.7	3.8	4.8	CHB
IS2	100	100	100	90	100	90	7.6	7.2	7.6	HCC
IS3	70	100	100	10	0	10	6.5	5.2	5.5	CHB
IS4	90	100	10	0	10	0	7.6	6.2	3.3	CHB
IS5	100	40	20	0	0	0	7.6	3.6	4.1	CHB
IS6	70	90	90	20	10	90	5.7	4.1	4.5	HCC
IS7	100	100	90	0	10	10	7.2	3.1	3.4	LC
IS8	80	100	60	0	0	60	7.6	4.0	4.5	CHB
IS9	100	100	10	0	0	0	6.0	4.5	4.8	HCC
SS1	0	60	0	80	0	80	7.6	4.2	5.4	HCC
SS2	80	100	90	10	90	10	6.6	7.6	5.9	ASC
SS3	100	90	60	10	0	0	6.5	4.3	2.8	ASC
SS6	30	100	10	0	0	10	3.9	4.4	4.1	CHB
SS7	80	100	100	0	0	0	7.6	2.8	2.6	ASC
SS8	0	100	90	0	20	0	7.6	5.4	3.6	CHB
SS9	0	80	20	0	10	0	7.6	4.0	2.6	CHB
SS10	50	20	40	0	0	40	7.3	3.9	2.6	CHB
SS11	100	100	100	0	0	0	6.1	6.3	3.8	CHB

IS: interferon induced seroconverter; SS: spontaneous seroconverter; ASC: asymptomatic carriers; CHB: chronic hepatitis B; LC: cirrhosis; HCC: hepatocellular carcinoma.

ratio of T1762A quasi-species between seroconverters and non-seroconverters, indicating that it might be a marker preceding seroconversion in HBV/genotype C-infected patients as reported previously [35–37].

HBeAg seroconversion is an incomplete marker of immune control, although most patients experience some clinical benefit from it [38,39]. Previous studies have shown that the average rate of spontaneous HBeAg seroconversion in patients with chronic hepatitis B is about 10% per year [40,41]. HBeAg seroconversion associated with incomplete viral suppression may result in the emergence of the *precore* mutant and attendant chronic sequelae. Mutations in the *precore* and *core* promoter regions of the HBV genome have been reported in many HBeAg-negative CHB patients. Longitudinal studies found that the A1896 mutation emerges or is selected around the time of HBeAg seroconversion, and high *precore* mutant ratios have been associated with persistent hepatitis after anti-HBe seroconversion [42]. Patients who continued to have high HBV DNA titres after HBe seroconversion had a lower genetic heterogeneity but more often had the *precore* mutant.

The limitations of this study were, the small size of study group, only 10 clones per sample, and a small region for analysis of genetic distance. In addition, the *X/precore/core* region is a highly conserved region, investigation of another region of the HBV genome, such as the polymerase, might help us to better understand the evolution of quasi-species of HBV.

In conclusion, the distinctly greater viral diversity after seroconversion in HBeAg seroconverters could be related to increased HBV-specific T-cell responses and escape mutants which arise from selective pressure caused by host immune activity. Long-term follow-up is required to determine whether hepatitis B viral diversity decreases or remains at a high level. Further study will be needed to elucidate the relationship between seroconversion and viral quasi-species in relation to antiviral therapy.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

References

[1] Domingo E, Holland JJ. RNA virus mutations and fitness for survival. Annu Rev Microbiol 1997;51:151–178.

[2] Arataki K, Kumada H, Toyota K, Ohishi W, Takahashi S, Tazuma S, et al. Evolution of hepatitis C virus quasispecies during ribavirin and interferon-alpha-2b combination therapy and interferon-alpha-2b monotherapy. Intervirology 2006;49:352–361.

[3] Laskus T, Wilkinson J, Gallegos-Orozco JF, Radkowski M, Adair DM, Nowicki M, et al. Analysis of hepatitis C virus quasispecies transmission and evolution in patients infected through blood transfusion. Gastroenterology 2004;127:764–776.

[4] Pawlotsky JM. Hepatitis C virus genetic variability: pathogenic and clinical implications. Clin Liver Dis 2003;7:45–66.

[5] Pellerin M, Lopez-Aguirre Y, Penin F, Dhumeaux D, Pawlotsky JM. Hepatitis C virus quasispecies variability modulates nonstructural protein 5A transcriptional activation, pointing to cellular compartmentalization of virus–host interactions. J Virol 2004;78:4617–4627.

[6] Fan X, Mao Q, Zhou D, Lu Y, Xing J, Xu Y, et al. High diversity of hepatitis C viral quasispecies is associated with early virological response in patients undergoing antiviral therapy. Hepatology 2009;50:1765–1772.

[7] da Silva J. The evolutionary adaptation of HIV-1 to specific immunity. Curr HIV Res 2003;1:363–371.

[8] Edwards CT, Holmes EC, Pybus OG, Wilson DJ, Viscidi RP, Abrams EJ, et al. Evolution of the human immunodeficiency virus envelope gene is dominated by purifying selection. Genetics 2006;174:1441–1453.

[9] Lipsitch M, O'Hagan JJ. Patterns of antigenic diversity and the mechanisms that maintain them. J R Soc Interface 2007;4:787–802.
[10] Ross HA, Rodrigo AG. Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration. J Virol 2002;76:11715–11720.
[11] Villet S, Pichoud C, Villeneuve JP, Trepo C, Zoulim F. Selection of a multiple drug-resistant hepatitis B virus strain in a liver-transplanted patient. Gastroenterology 2006;131:1253–1261.
[12] Yim HJ, Hussain M, Liu Y, Wong SN, Fung SK, Lok AS. Evolution of multi-drug resistant hepatitis B virus during sequential therapy. Hepatology 2006;44:703–712.
[13] Bertoletti A, Maini M, Williams R. Role of hepatitis B virus specific cytotoxic T cells in liver damage and viral control. Antiviral Res 2003;60:61–66.
[14] Huang CF, Lin SS, Ho YC, Chen FL, Yang CC. The immune response induced by hepatitis B virus principal antigens. Cell Mol Immunol 2006;3:97–106.
[15] Panther E, Spangenberg HC, Neumann-Haefelin C, Rosler K, Blum HE, von Weizsacker F, et al. The role of the virus specific T-cell response in acute and chronic HBV and HCV infection. Z Gastroenterol 2004;42:39–46.
[16] Tan AT, Koh S, Goh V, Bertoletti A. Understanding the immunopathogenesis of chronic hepatitis B virus: an Asian prospective. J Gastroenterol Hepatol 2008;23:833–843.
[17] Bruss V, Gerlich WH. Formation of transmembraneous hepatitis B e-antigen by cotranslational in vitro processing of the viral precore protein. Virology 1988;163:268–275.
[18] Garcia PD, Ou JH, Rutter WJ, Walter P. Targeting of the hepatitis B virus precore protein to the endoplasmic reticulum membrane: after signal peptide cleavage translocation can be aborted and the product released into the cytoplasm. J Cell Biol 1988;106:1093–1104.
[19] Carman WF, Jacyna MR, Hadziyannis S, Karayiannis P, McGarvey MJ, Makris A, et al. Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet 1989;2:588–591.
[20] Okamoto H, Tsuda F, Akahane Y, Sugai Y, Yoshiba M, Moriyama K, et al. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. J Virol 1994;68:8102–8110.
[21] Usuda S, Okamoto H, Iwanari H, Baba K, Tsuda F, Miyakawa Y, et al. Serological detection of hepatitis B virus genotypes by ELISA with monoclonal antibodies to type-specific epitopes in the preS2-region product. J Virol Meth 1999;80:97–112.
[22] Simmonds P. Recombination and selection in the evolution of picornaviruses and other Mammalian positive-stranded RNA viruses. J Virol 2006;80: 11124–11140.
[23] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24: 1596–1599.
[24] Osiowy C, Giles E, Tanaka Y, Mizokami M, Minuk GY. Molecular evolution of hepatitis B virus over 25 years. J Virol 2006;80:10307–10314.
[25] Bozkaya H, Ayola B, Lok AS. High rate of mutations in the hepatitis B core gene during the immune clearance phase of chronic hepatitis B virus infection. Hepatology 1996;24:32–37.
[26] Hannoun C, Horal P, Lindh M. Long-term mutation rates in the hepatitis B virus genome. J Gen Virol 2000;81:75–83.
[27] Simmonds P. The origin and evolution of hepatitis viruses in humans. J Gen Virol 2001;82:693–712.
[28] Fares MA, Holmes EC. A revised evolutionary history of hepatitis B virus (HBV). J Mol Evol 2002;54:807–814.
[29] Locarnini S. Molecular virology and the development of resistant mutants: implications for therapy. Semin Liver Dis 2005;25 (Suppl. 1):9–19.
[30] Stumpf MP, Pybus OG. Genetic diversity and models of viral evolution for the hepatitis C virus. FEMS Microbiol Lett 2002;214:143–152.
[31] Lim SG, Cheng Y, Guindon S, Seet BL, Lee LY, Hu P, et al. Viral quasi-species evolution during hepatitis Be antigen seroconversion. Gastroenterology 2007;133:951–958.
[32] Chan HL, Hussain M, Lok AS. Different hepatitis B virus genotypes are associated with different mutations in the core promoter and precore regions during hepatitis B e antigen seroconversion. Hepatology 1999;29:976–984.
[33] Lindh M, Hannoun C, Dhillon AP, Norkrans G, Horal P. Core promoter mutations and genotypes in relation to viral replication and liver damage in East Asian hepatitis B virus carriers. J Infect Dis 1999;179:775–782.
[34] Blackberg J, Kidd-Ljunggren K. Genotypic differences in the hepatitis B virus core promoter and precore sequences during seroconversion from HBeAg to anti-HBe. J Med Virol 2000;60:107–112.
[35] Lindh M, Gustavson C, Mardberg K, Norkrans G, Dhillon AP, Horal P. Mutation of nucleotide 1, 762 in the core promoter region during hepatitis B e seroconversion and its relation to liver damage in hepatitis B e antigen carriers. J Med Virol 1998;55:185–190.
[36] Kim YS, Kim SI, Hwang SG, Kim JO, Cho JY, Lee JS, et al. Diversity of core promoter mutations in immune clearance phase of chronic HBV infection. Eur J Gastroenterol Hepatol 1999;11:821–825.
[37] Liu CJ, Chen PJ, Lai MY, Kao JH, Chen DS. Evolution of precore/core promoter mutations in hepatitis B carriers with hepatitis B e antigen seroreversion. J Med Virol 2004;74:237–245.
[38] Hsu YS, Chien RN, Yeh CT, Sheen IS, Chiou HY, Chu CM, et al. Long-term outcome after spontaneous HBeAg seroconversion in patients with chronic hepatitis B. Hepatology 2002;35:1522–1527.
[39] Hui CK, Leung N, Shek TW, Yao H, Lee WK, Lai JY, et al. Sustained disease remission after spontaneous HBeAg seroconversion is associated with reduction in fibrosis progression in chronic hepatitis B Chinese patients. Hepatology 2007;46:690–698.
[40] Liaw YF, Chu CM, Su IJ, Huang MJ, Lin DY, Chang-Chien CS. Clinical and histological events preceding hepatitis B e antigen seroconversion in chronic type B hepatitis. Gastroenterology 1983;84:216–219.
[41] Lok AS, Lai CL, Wu PC, Leung EK, Lam TS. Spontaneous hepatitis B e antigen to antibody seroconversion and reversion in Chinese patients with chronic hepatitis B virus infection. Gastroenterology 1987;92:1839–1843.
[42] Chu CM, Yeh CT, Lee CS, Sheen IS, Liaw YF. Precore stop mutant in HBeAg-positive patients with chronic hepatitis B: clinical characteristics and correlation with the course of HBeAg-to-anti-HBe seroconversion. J Clin Microbiol 2002;40:16–21.

Network based analysis of hepatitis C virus Core and NS4B protein interactions†

Lokesh P. Tripathi,[a] Chikako Kataoka,[b] Shuhei Taguwa,[b] Kohji Moriishi,[b] Yoshio Mori,[b] Yoshiharu Matsuura[b] and Kenji Mizuguchi*[a]

Received 15th July 2010, Accepted 21st September 2010
DOI: 10.1039/c0mb00103a

Hepatitis C virus (HCV) is a major cause of chronic liver disease worldwide. Here we attempt to further our understanding of the biological context of protein interactions in HCV pathogenesis, by investigating interactions between HCV proteins Core and NS4B and human host proteins. Using the yeast two-hybrid (Y2H) membrane protein system, eleven human host proteins interacting with Core and 45 interacting with NS4B were identified, most of which are novel. These interactions were used to infer overall protein interaction maps linking the viral proteins with components of the host cellular networks. Core and NS4B proteins contribute to highly compact interaction networks that may enable the virus to respond rapidly to host physiological responses to HCV infection. Analysis of the interaction networks highlighted enriched biological pathways likely influenced in HCV infection. Inspection of individual interactions offered further insights into the possible mechanisms that permit HCV to evade the host immune response and appropriate host metabolic machinery. Follow-up cellular assays with cell lines infected with HCV genotype 1b and 2a strains validated Core interacting proteins ENO1 and SLC25A5 and host protein PXN as novel regulators of HCV replication and viral production. ENO1 siRNA knockdown was found to inhibit HCV replication in both the HCV genotypes and viral RNA release in genotype 2a. PXN siRNA inhibition was observed to inhibit replication specifically in genotype 1b but not in genotype 2a, while SLC25A5 siRNA facilitated a minor increase in the viral RNA release in genotype 2a. Thus, our analysis can provide potential targets for more effective anti-HCV therapeutic intervention.

1. Introduction

Hepatitis C virus (HCV) is the causative agent of chronic liver disease including liver steatosis, cirrhosis and hepatocellular carcinoma (HCC) and infects nearly 3% of the population worldwide. HCV is a positive single-stranded RNA virus with a single 9600 nucleotide ORF flanked by 5′ and 3′-UTRs. The HCV ORF encodes a 3000 amino acid polyprotein, which undergoes proteolytic processing by host and viral proteases to yield four structural (Core, E1, E2 and p7) and six non-structural (NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins.[1–3] HCV variants span six genotypes that display phylogenetic heterogeneity, differences in infectivity and interferon sensitivity.[4] However, despite a wealth of concerted research, a precise understanding of the molecular mechanisms underlying HCV pathology remains elusive.

Most genes and proteins function in a complex web of interactions. Thus, the study of protein–protein interactions (PPIs) is critical to understanding the cellular networks that regulate the physiology of a living organism. The increasing availability of PPI data for human and host–pathogen interactions has led to increasing efforts in understanding the network basis of human diseases and pathogenesis.[5,6] In particular, the increasing availability of large scale interaction data between viral and human host proteins is likely to lead to a better understanding of viral pathogenesis and help identify novel targets for experimental and therapeutic intervention.[7–9] Comprehensive analyses of yeast two-hybrid (Y2H) screens have been employed to investigate the interactions of HCV,[7] Epstein–Barr virus[10] herpesviral[11] proteins with host factors. Analysis of such interactions suggests that viral (and bacterial) pathogens preferably interact with host proteins either engaged in a large number of interactions or critical to the integrity of the host cellular networks.[7,10]

Here, we report the host biological processes likely to be perturbed by HCV Core and NS4B proteins by virtue of inferred PPI networks. Core, also known as capsid protein, is spliced from the polyprotein by signal peptidase and further processed into a highly conserved 21 kDa mature form by the signal peptide peptidase; this processing facilitates its transfer to the detergent-resistant membrane fraction where virus replication and assembly take place.[12] Core also localises to the nucleus, which is essential for efficient viral propagation and development of HCV pathogenicity.[3,13] Core is a multifunctional protein implicated in RNA binding and as a pathogenic factor, which induces steatosis and HCC and thus, a promising target for anti-HCV therapy.[14,15] NS4B, the least

[a] *National Institute of Biomedical Innovation, 7-6-8 Asagi-Saito, Ibaraki-City, Osaka, 567-0085, Japan. E-mail: kenji@nibio.go.jp*
[b] *Department of Molecular Virology, Research Institute for Microbial Diseases, Osaka University, Osaka, 565-0871, Japan*

† Electronic supplementary information (ESI) available: PPI networks analysed in the study and their functional associations. See DOI: 10.1039/c0mb00103a

characterised HCV protein, is a 27 kDa non-structural integral membrane protein located in the ER membrane, which induces membrane changes and facilitates HCV replication in the host cells, though recent reports suggest that it may function in HCV pathogenesis and viral assembly.[14,16] Since Core and NS4B proteins are primarily associated with the ER membrane, they were employed as baits to screen against a library of human cDNAs using the Y2H membrane protein approach, which identifies PPIs involving integral membrane proteins and membrane-associated proteins in an *in vivo* setting.[17] We identified 11 interactions for Core and 45 interactions for NS4B, nearly all of which are previously uncharacterised. By extending these interactions to include human protein interaction data, our analysis provided insights into the functional pathways likely to be associated with HCV–host interactions in HCV pathogenesis, a better understanding of which may help identify new targets for anti-HCV therapeutic intervention.

2. Results and discussion

2.1 Identifying host proteins that interact with HCV Core and NS4B proteins

Since Core and NS4B are primarily localised to the ER membrane, to investigate their biological associations, we performed a series of Y2H screens customised for characterising the PPIs involving integral membrane and membrane-associated proteins (see Materials and Methods). Analysis of positive colonies revealed 11 interactors for Core protein and 45 interactors for NS4B protein (Table 1).

Nine of the 11 host proteins interacting with Core (Table 1) are novel findings but the other two interactions are known; signal peptide peptidase (HM13), an ER localised protein, is crucial for the intramembrane processing of the Core protein, facilitating its localisation and viral propagation;[12] proteasome subunit alpha type 7 (PSMA7) is involved in regulating HCV internal ribosome entry site (IRES), which is essential for HCV replication.[18] These results suggest that the PPIs detected by our screening approach may closely reflect Core interactions *in vivo*. Among the other interacting proteins, four localise to mitochondria and are likely involved in oxidative electron transfer (ETFB; NDUFS2) and solute transport (SLC25A5; TOMM20), which may be a consequence of known Core localisation to the mitochondrial outer membrane.[19] Additionally, Core interacting proteins Alpha Enolase (ENO1), Ferritin light chain (FTL) and SLC25A5 are perturbed in cancerous tissues from HCC patients with HCV infection.[20–22] These observations suggest potential roles for the above-identified Core protein interactions in HCV infection.

NS4B protein was found to interact with 45 host proteins (Table 1), nearly all of which are novel interactions. A significant proportion of these mapped to either the membrane component (GO:0016020; 17 of 45, 38%; $p = 0.04243$), or the extracellular region (GO:0005576; 12 of 45, 27%; $p = 3.64 \times 10^{-4}$), while five (APOA1, APOB, F2, FGG, LRG1) localise to both compartments. The NS4B interactions with a large number of host proteins (especially membrane proteins) may be crucial to its ability to induce membrane alterations termed membranous webs (MW), which host the HCV replication complex. It appears consistent with the suggested role of NS4B protein as an important hub in the virus–host interaction network.[16,23]

The absence of overlap between the PPIs identified in our approach and a previous large scale study[7] may be attributed to differences in screening approaches and experimental settings. Since our approach seeks to investigate the interactions associated with Core and NS4B in their membrane setting, it is more likely to fish out associations that may not be easily detected using the standard Y2H screening and co-immunoprecipitation assays such as those employed by de Chassey *et al.*[7] This situation would be especially true for NS4B, which unlike Core is not detected outside the membrane fraction, thus explaining the 45 interactions for NS4B reported by our approach compared to the one reported previously.[7] Our observations also highlight the significance of employing specific approaches to investigating different aspects of host–pathogen interactions in general.

2.2 Topological analysis of Core and NS4B protein interaction networks

To further understand the biological processes likely targeted by HCV, we expanded the Y2H-derived interactions by incorporating the secondary interactors of the human proteins that interact with the Core and NS4B proteins to derive extended PPI networks (Fig. 1a and b). The Core extended PPI network was made up of 208 entities (genes) with 1063 interactions between them (Table S1–S3, ESI†). For comparison, we also derived an extended PPI network for Core interactions reported by de Chassey *et al.*[7] (Table S4, ESI†). The NS4B extended PPI network was made up of 253 entities (genes) with 481 interactions between them (Table S1–S3, ESI†). First, we computed *node degree distribution* and *characteristic/average path length* measures to capture the topology of the Core and NS4B extended PPI networks (Fig. 3A and B). The degree of a protein, which corresponds to the number of its interacting partners, provides some insights into its biological relevance, since a higher degree may likely correspond to a higher ability to influence biological networks. It is also a useful measure to distinguish real world and random networks. In most interactome networks, a few nodes called "hubs" have a high degree and most nodes have a low degree, while in random networks the degree is uniformly distributed. Average path lengths provide an approximate measure of the relative ease and speed of dissemination of signalling information among network components.

Our analysis revealed that the average degree of the Core membrane protein yeast two-hybrid (MY2H) network (9.75) is on par with the human interactome (9.3), though shorter than the average degree estimated for the Core de Chassey extended network (14.7) (Table S4, ESI†). The Core MY2H network has a shorter characteristic path length vis-à-vis the human interactome (2.9 *versus* 4.04) and on par with that of Core de Chassey network (2.97), which is consistent with previous observations on the HCV protein infection network.[7] While the average degree of the NS4B MY2H network (3.4) is substantially lower than that of the

Table 1 List of host proteins interacting with HCV Core and NS4B proteins, identified by Y2H screens

List of host proteins interacting with the Core protein		
Gene ID	Official symbol	Description
1937	EEF1G	Eukaryotic translation elongation factor 1 gamma
1964	EIF1AX	Eukaryotic translation initiation factor 1A, X-linked
2023	ENO1	Enolase 1 (alpha)
2109	EFTB	Electron-transfer-flavoprotein, beta polypeptide
2512	FTL	Ferritin, light polypeptide
292	SLC25A5	Solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5
4720	NDUFS2	NADH dehydrogenase (ubiquinone) Fe–S protein 2, 49 kDa (NADH-coenzyme Q reductase)
5265	SERPINA1	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
5688	PSMA7	Proteasome (prosome, macropain) subunit, alpha type, 7
81502	HM13	Histocompatibility (minor) 13
9804	TOMM20	Translocase of outer mitochondrial membrane 20 homolog (yeast)
List of host proteins interacting with the HCV NS4B protein		
Gene ID	Official symbol	Description
10130	PDIA6	Protein disulfide isomerase family A, member 6
10682	EBP	Emopamil binding protein (sterol isomerase)
116844	LRG1	Leucine-rich alpha-2-glycoprotein 1
1209	CLPTM1	Cleft lip and palate associated transmembrane protein 1
132299	OCIAD2	OCIA domain containing 2
1528	CYB5A	Cytochrome b5 type A (microsomal)
154467	C6orf129	Chromosome 6 open reading frame 129
1571	CYP2E1	Cytochrome P450, family 2, subfamily E, polypeptide 1
196410	METTL7B	Methyltransferase like 7B
200185	KRTCAP2	Keratinocyte associated protein 2
2013	EMP2	Epithelial membrane protein 2
2147	F2	Coagulation factor II (thrombin)
2220	FCN2	Ficolin (collagen/fibrinogen domain containing lectin) 2 (hucolin)
2266	FGG	Fibrinogen gamma chain
2267	FGL1	Fibrinogen-like 1
27173	SLC39A1	Solute carrier family 39 (zinc transporter), member 1
2731	GLDC	Glycine dehydrogenase (decarboxylating)
286451	YIPF6	Yip1 domain family, member 6
334	APLP2	Amyloid beta (A4) precursor-like protein 2
335	APOA1	Apolipoprotein A-I
338	APOB	Apolipoprotein B (including Ag(x) antigen)
3732	CD82	CD82 molecule
4267	CD99	CD99 molecule
4513	COX2	Cytochrome c oxidase subunit II
4538	ND4	NADH dehydrogenase, subunit 4 (complex I)
4924	NUCB1	Nucleobindin 1
51075	TMX2	Thioredoxin-related transmembrane protein 2
51643	TMBIM4	Transmembrane BAX inhibitor motif containing 4
517	ATP5G2	ATP synthase, H^+ transporting, mitochondrial F0 complex, subunit C2 (subunit 9)
5265	SERPINA1	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
5355	PLP2	Proteolipid protein 2 (colonic epithelium-enriched)
5446	PON3	Paraoxonase 3
54657	UGT1A4	UDP glucuronosyltransferase 1 family, polypeptide A4
54658	UGT1A1	UDP glucuronosyltransferase 1 family, polypeptide A1
5479	PPIB	Peptidylprolyl isomerase B (cyclophilin B)
563	AZGP1	Alpha-2-glycoprotein 1, zinc-binding
56851	C15orf24	Chromosome 15 open reading frame 24
57817	HAMP	Hepcidin antimicrobial peptide
5950	RBP4	Retinol binding protein 4, plasma
6048	RNF5	Ring finger protein 5
6522	SLC4A2	Solute carrier family 4, anion exchanger, member 2 (erythrocyte membrane protein band 3-like 1)
7905	REEP5	Receptor accessory protein 5
84975	MFSD5	Major facilitator superfamily domain containing 5
9204	ZMYM6	Zinc finger, MYM-type 6
967	CD63	CD63 molecule

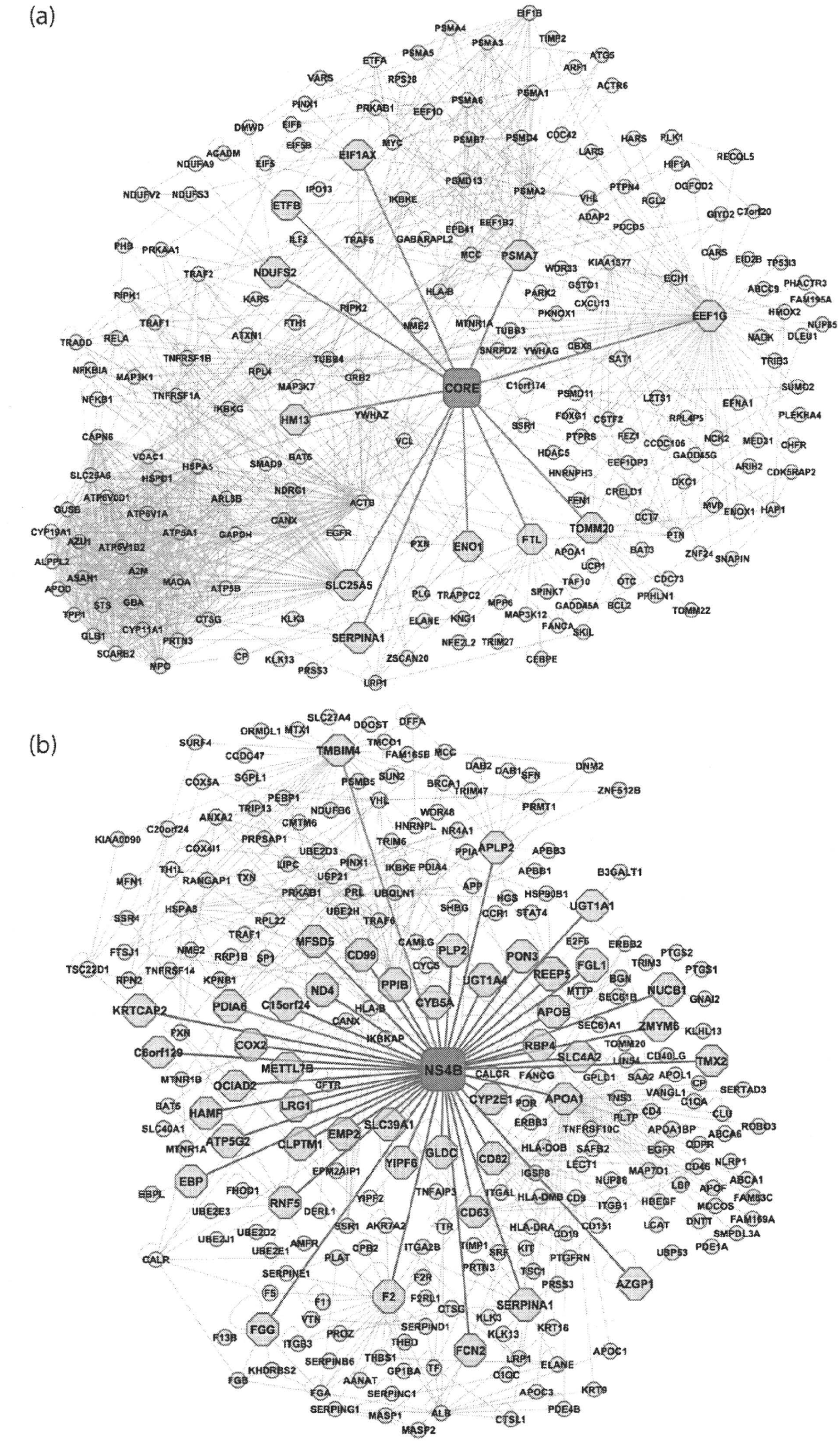

Fig. 1 Graphical representation of HCV (a) Core protein and (b) NS4B protein extended protein–protein interaction (PPI) networks. Red node: Core and NS4B protein; blue edge: Core and NS4B yeast two-hybrid (Y2H) interactions; green node: host proteins identified as interacting partners of Core and NS4B by Y2H membrane protein system; pink node: secondary interactors of the host proteins interacting with Core and NS4B; grey edge: interactions between human proteins. The node sizes differ for better clarity and do not reflect any topological attributes.

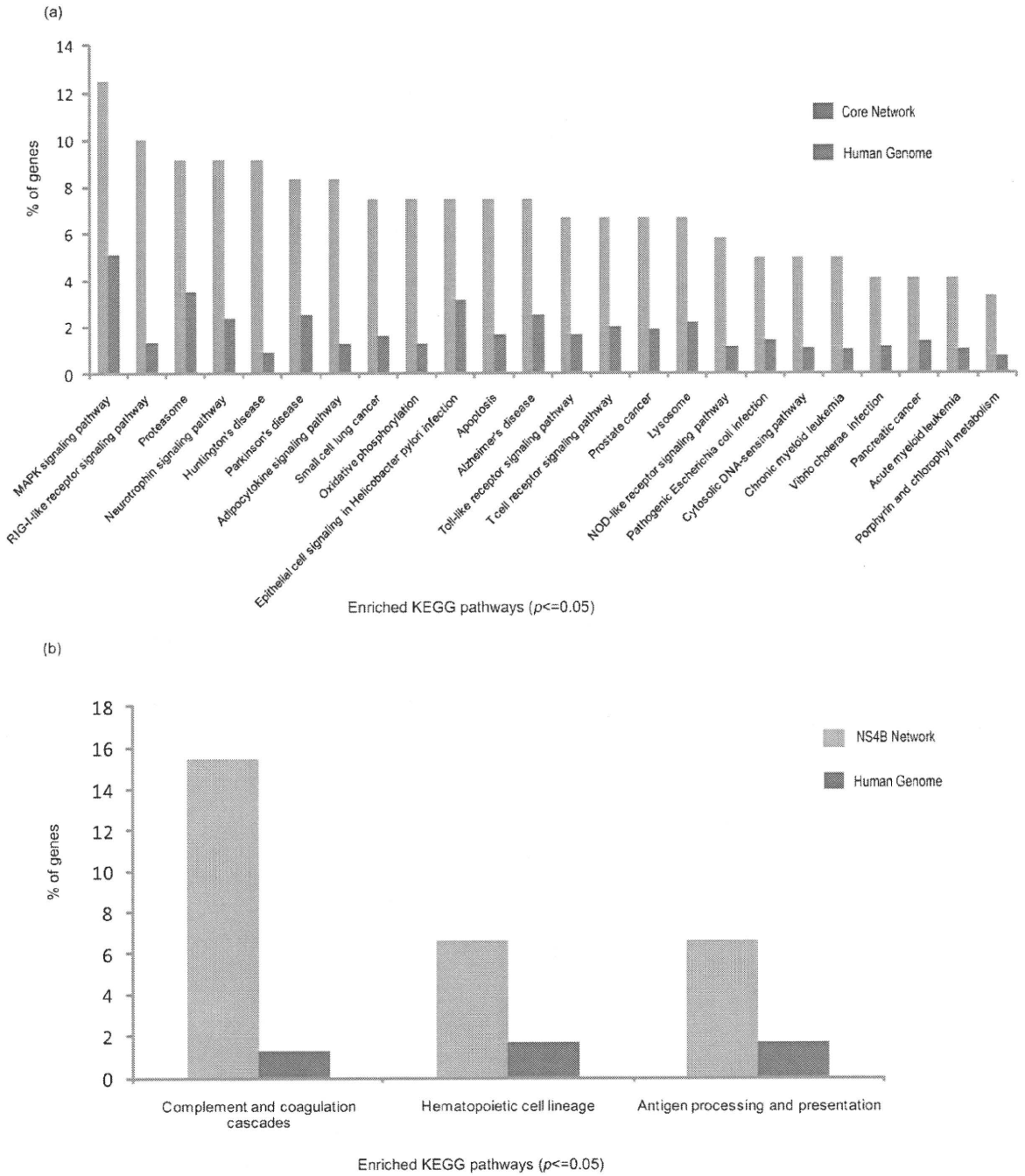

Fig. 2 KEGG pathway associations enriched in (a) Core and (b) NS4B extended PPI networks. The pathway labels are mapped to *x*-axis while the *y*-axis represents the % of genes mapped to a given KEGG pathway within the network and in the human genome.

Core MY2H network (9.75) and the human interactome (9.3), the characteristic path length is on par with the former (3.3 *versus* 2.9) and shorter than the human interactome (3.3 *versus* 4.04). Our observations suggest that the compactness of Core and NS4B interaction networks may facilitate rapid propagation of signaling information and allow the virus to respond rapidly to host mobilisation against HCV infection.

2.3 Functional analysis of Core and NS4B interactions with host proteins

Next, we investigated the extended networks for enrichment of specific biological associations (KEGG pathways,[24] OMIM phenotypes[25] and Gene Ontology [GO][26] terms). The analysis revealed that Core and NS4B protein networks were enriched in largely non-overlapping functional associations, indicating

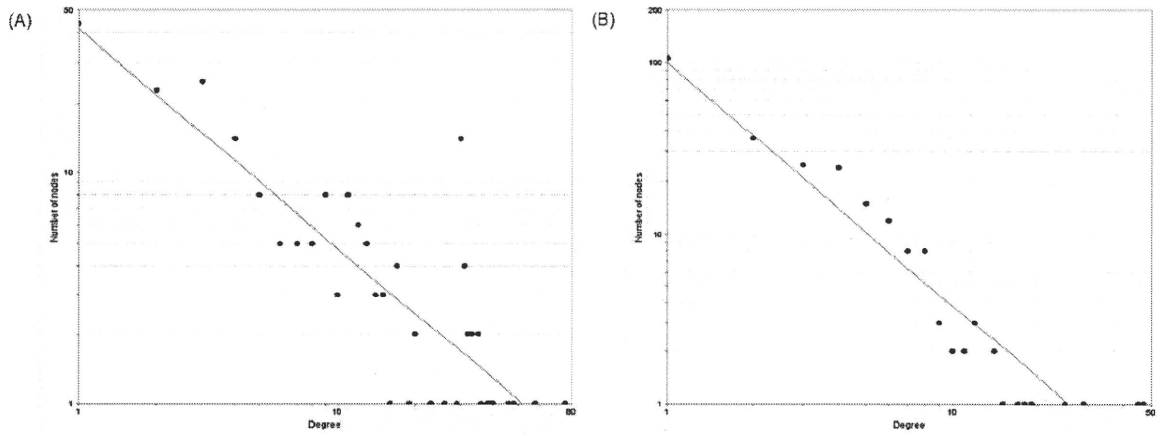

Fig. 3 Graphical representation of node degree distribution for Core and NS4B extended PPI networks. The node degree k is represented on the x-axis while the number of nodes mapped to a specific degree is represented on y-axis. The inverse trend between degree distribution and the number of proteins indicates a non-random network. The graphs correspond to (A) Core extended PPI network (R-squared value 0.727) and (B) NS4B extended network (R-squared value 0.895).

that they occupy different niches in HCV infection (Table 2; Fig. 2). Below, we first describe our observations on the Core network followed by the NS4B network.

2.3.1 Functional analysis of Core interaction network. The analysis of the extended Core interaction network revealed an enrichment of 24 KEGG pathways ($p \leq 0.05$) (Table 2), encompassing host immune response and oxidative and non-oxidative metabolism, which are likely to be significantly affected by HCV infection. Seventeen of the 24 enriched associations overlapped with the enriched KEGG pathway associations inferred for the Core de Chassey network (Table S4, ESI†). Furthermore, to assess the robustness of enriched KEGG pathway associations, we explored the overlap of PPIs in the Core network with those documented in the I2D database.[27] We observed that 559 of the 1052 secondary interactions (see Material and Methods) in the Core interaction network were represented in I2D; the genes associated with these interactions were mapped to 22 enriched KEGG pathways, of which 20 were shared with the enriched KEGG pathway associations for the Core interaction network (data not shown).

Table 2 The number of proteins mapped to KEGG pathways significantly enriched ($p \leq 0.05$) in Core and NS4B extended PPI networks

KEGG pathway	Core network	NS4B network
Complement and coagulation cascades	—	21
MAPK signaling pathway	15	—
RIG-I-like receptor signaling pathway	12	—
Huntington's disease	11	—
Neurotrophin signaling pathway	11	—
Proteasome	10	—
Adipocytokine signaling pathway	10	—
Parkinson's disease	9	—
Alzheimer's disease	9	—
Antigen processing and presentation	—	9
Apoptosis	9	—
Epithelial cell signaling in *Helicobacter pylori* infection	9	—
Hematopoietic cell lineage	—	9
Oxidative phosphorylation	9	—
Small cell lung cancer	9	—
Lysosome	8	—
Prostate cancer	8	—
T cell receptor signaling pathway	8	—
Toll-like receptor signaling pathway	8	—
NOD-like receptor signaling pathway	7	—
Chronic myeloid leukemia	6	—
Cytosolic DNA-sensing pathway	6	—
Pathogenic *Escherichia coli* infection	6	—
Acute myeloid leukemia	5	—
Pancreatic cancer	5	—
Vibrio cholerae infection	5	—
Porphyrin and chlorophyll metabolism	4	—

2.3.1.1 Immune response. Immune response to HCV infection includes the recognition of the HCV RNA and proteins as pathogen associated molecular patterns (PAMPs) by macrophages and dendritic cells expressing germline-encoded pattern-recognition receptors (PRRs) such as Toll-like receptors (TLRs) and RIG-I-like receptors (RLRs), a family of RNA helicases. These events induce the production of Type I interferons (IFN-α/β) and inflammatory cytokines in the infected hepatocytes, which then activate downstream processes such as T-cell signalling for viral clearance.[28–32] The persistence of HCV in the host is attributed to its ability to hinder and evade the host immune response, which is regulated by the interplay between the HCV proteins and the components of the host immune system.[31,33,34]

Analysis of the Core interaction network revealed that six of the 11 Core interacting proteins interact with proteins mapped to one or more of the following KEGG pathways: "RIG-I-like receptor signaling" (12 of 208, $p = 4.0117 \times 10^{-7}$), "T cell receptor signaling" (8 of 208, $p = 0.0072$), "NOD-like receptor signaling" (7 of 208, $p = 0.00267$), "Toll-like receptor signaling" (8 of 208, $p = 0.0054$) and "Adipocytokine signaling" (10 of 208, $p = 1.637 \times 10^{-5}$). All of these pathways are related to innate and adaptive immunity (Fig. 1a and 4a, Table 2, Table S3, ESI†). Thus, Core protein may partly cause perturbations in the host immune response by virtue of these interactions. Several genes involved in these processes are implicated in

HCV infection. For instance, hepatic RELA mRNA levels are suppressed in chronic hepatitis C patients and associated with increased liver fibrosis,[35] while RELA and TRAF2 knockdowns with siRNA resulted in a substantial reduction in HCV replication.[36] Additionally, the interaction of HCV protein NS5A with TRAF2 inhibits NFKB activation and thereby disrupts the host immune response,[37] which also suggests probable links between HCV replication and NFKB activation.[36] The expression of NFKBIA, a critical regulator of NFKB activation, is known to be increased by HCV Core protein, resulting in suppression of pro-inflammatory genes downstream of NFKB.[38] IKBKG (IKK Gamma), an anti-apoptotic protein, is essential for NFKB activation and modulates tumour necrosis factor (TNF)-mediated apoptosis.[39] IKBKG mutations are associated with immune deficiency phenotype (Table S5, ESI†) and disruption of IKBKG activity may contribute to impaired immune response in HCV infection.

Some genes (PRKAA1 and PRKAB1) associated with the pro-inflammatory "Adipocytokine signaling pathway" also function in "Insulin signaling pathway", the disruption of which may contribute to insulin resistance (IR). IR is commonly observed in HCV infection and is associated with steatosis and fibrosis progression and impaired response to interferon-α anti-HCV therapy[40,41] and overexpression of HCV Core protein can induce IR in transgenic mice.[42] PRKAA1 expression is implicated in HCV infection,[43] suggesting that it may function in Core-induced IR.

The bulk of these genes (RELA, NFKB1, IKBKG, NFKBIA, TRAF2, PRKAA1) interact with SLC25A5 (Fig. 4a), suggesting that SLC25A5 may play an important role in Core perturbation of host innate immune response and liver fibrosis and possibly HCV replication.

2.3.1.2 Oxidative stress. HCV and other pathogens have evolved mechanisms to modulate host metabolism to facilitate their survival and propagation. ER stress, oxidative stress and mitochondrial dysfunction are some characteristic features associated with chronic hepatitis C infection.[44–46] Overexpression of HCV Core, NS3 and NS5A proteins is associated with increased production of reactive oxygen species (ROS), disrupted mitochondrial electron transport and altered Ca^{2+} homeostasis leading to perturbed Cytochrome c release from the mitochondria. This reaction together with induced insulin resistance eventually leads to accelerated fibrosis, HCC and DNA damage, while ensuring cell survival.[1,46–48]

Our Y2H screening identified NDUFS2, a mitochondrial protein essential for NADH to ubiquinone electron transfer, and ETFB, a mitochondrial electron-transfer flavoprotein as interacting partners of Core protein (Table 1). These interactions may permit Core protein to perturb oxidative electron transfer and consequently induce mitochondrial aberrations, which would be consistent with oxidative modification of mitochondrial respiratory complexes in pathogen infection.[49] Mutations in NDUFS2 (KEGG pathway "Oxidative phosphorylation" (10 of 208, $p = 0.0072$)) are associated with the OMIM phenotype "Mitochondrial complex I deficiency" (Table S5, ESI†), which causes several clinical disorders including liver disease. ETFB and its interacting partner ETFA are involved in beta-oxidation of fatty acids; ETFA displays a decreased activity during HCV replication, possibly contributing to steatosis,[50] suggesting that its interaction with ETFB may permit Core protein to interfere with mitochondrial fatty acid metabolism. This interference may contribute to lipid accumulation in hepatocytes, which facilitates viral entry, replication and assembly, insulin resistance and steatosis.[40,51–53]

In addition, Core may also induce mitochondrial perturbations *via* SLC25A5 protein which interacts with the subunits of ATP synthase (ATP5A1, ATP5B) and ATPase (ATP6V1A, ATP6V1B2, ATP6V0D1) enzymes. Interestingly, SLC25A5 (ANT2) in association with ATP synthase is involved in glycolytic ATP import into mitochondria, thus facilitating a shift to almost exclusively glycolytic metabolism in cancer cells.[54] Thus, Core interactions with SLC25A5 may provide an important link between oxidative stress and a shift in energy metabolism in HCV infection (see below).

2.3.1.3 Host energy metabolism and cell adhesion. HCV induction of oxidative stress is accompanied by a shift towards non-oxidative glucose metabolism to facilitate viral growth and is often characterised by elevated levels of glycolytic enzymes in the infected cells.[22,55] Our Y2H screening identified a novel interaction between Core protein and Alpha Enolase (ENO1), a key enzyme in the glycolytic pathway implicated in several disorders including metastatic cancer.[56] ENO1 was upregulated in response to HCV infection[22] and may be a key regulator in the shift towards glycolytic metabolism and viral replication. However, there is little understanding of the role of ENO1 in HCV infection and to the best of our knowledge no physical interaction between HCV proteins and ENO1 had been reported earlier.

Our network analysis revealed that some proteins interacting with ENO1 (ACTB, PXN) mapped to KEGG "Focal adhesion" pathway (Fig. 4b; Table 1, Table S3, ESI†). Focal adhesion regulates cell migration and its deregulation is linked to tumour progression and probably HCV propagation in the host.[7] Paxillin (PXN) is involved in cytoskeleton remodelling and a disruption in its activity by human papilloma virus (HPV) E6 protein is an important aspect of HPV pathogenesis.[57] Thus, the Core interaction with ENO1 may be important in HCV mediated cytoskeleton remodelling to facilitate viral propagation. SLC25A5 (ANT2), also implicated in cancer cell glycolysis,[54] interacts with proteins (ACTB, EGFR, PXN, VCL) involved in "Focal adhesion" (Fig. 4b; Table 1, Table S3, ESI†). Most of these interactions are associated with reasonable confidence levels.[58,59] Enhancement of EGFR signalling by HCV NS3/4A and NS5A proteins is important for viral replication and persistence.[60,61] Interestingly, EGFR is upregulated in lymph node metastasis in HCC, while SLC25A5 levels are downregulated,[20] suggesting that the Core interaction with SLC25A5 may be involved in regulating EGFR activity and host energy metabolism and consequently HCV replication and propagation.

HCV infection is also characterised by elevated hepatic iron levels (induced by ROC), which contributes to abnormalities in glucose metabolism and induction of insulin resistance, eventually leading to fibrosis.[62,63] Our Y2H screening identified interaction between Core and FTL, a subunit of the

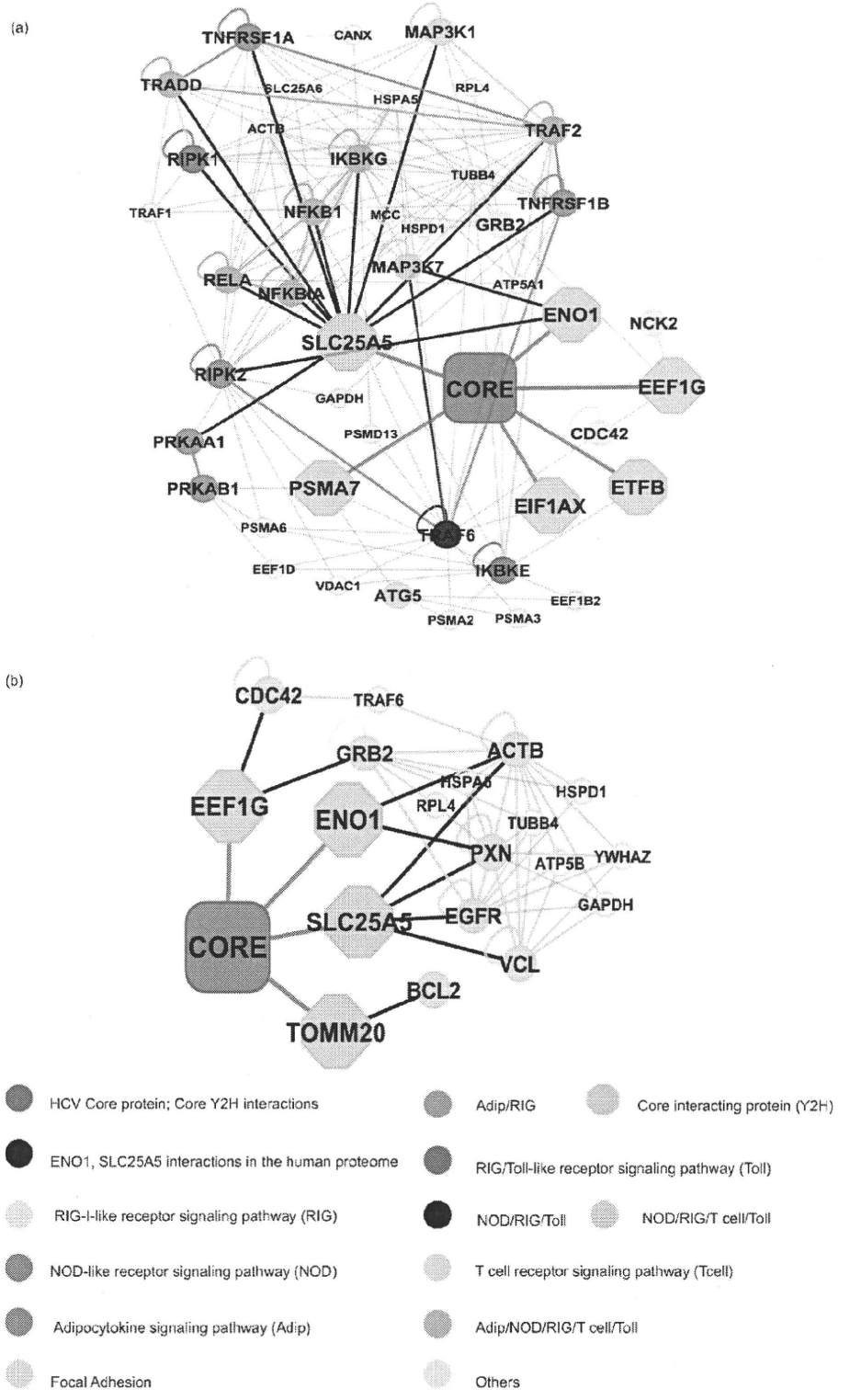

Fig. 4 Functional associations of the Core network. (a) Network illustration of interactions between HCV Core interacting proteins and host proteins mapped to five KEGG pathways— "RIG-I-like receptor signaling", "T cell receptor signaling", "NOD-like receptor signaling", "Toll-like receptor signaling" and "Adipocytokine signaling". SLC25A5 and ENO1 interactions are highlighted in black. (b) Network illustration of interactions between HCV Core interacting proteins and host proteins mapped to KEGG "Focal adhesion" pathway. The node sizes differ for better clarity and do not reflect any topological attributes.

intracellular iron storage protein Ferritin, which may facilitate Core induced disruptions in host iron metabolism.

2.3.2 Functional analysis of NS4B interaction network. To further understand the biological context of NS4B interactions, we examined the overrepresentation of specific biological associations within the NS4B PPI network. The analysis revealed an enrichment of three KEGG pathways ($p \leq 0.05$) "Complement and coagulation cascades", "Hematopoietic cell lineage" and "Oxidative phosphorylation" likely to be significantly affected by NS4B interactions in HCV infection (Table 2). To assess the robustness of enriched KEGG pathway associations, we explored the overlap of PPIs in the NS4B network with those documented in the I2D database.[27] We observed that 396 of the 436 secondary interactions (see Materials and Methods) in the NS4B interaction network were represented in I2D;[27] the genes associated with these interactions were mapped to two enriched KEGG pathways ("Complement and coagulation cascades" and "Hematopoietic cell lineage"; data not shown), both of which were among the enriched KEGG pathway associations for the NS4B interaction network.

2.3.2.1 Complement and coagulation. A significant number of proteins in the NS4B network mapped to KEGG pathway "Complement and coagulation cascades" (21 of 254, 5%; $p = 2.49 \times 10^{-14}$; Fig. 5, Table 2, Table S3, ESI†), which functions in the host innate immune response against pathogen invasion and clearance of viral antigens from the blood of the infected hosts. Overexpression of HCV Core protein activates the coagulation pathway *via* hepatic inflammation, which contributes to fibrosis, disruption of host T-cell mediated

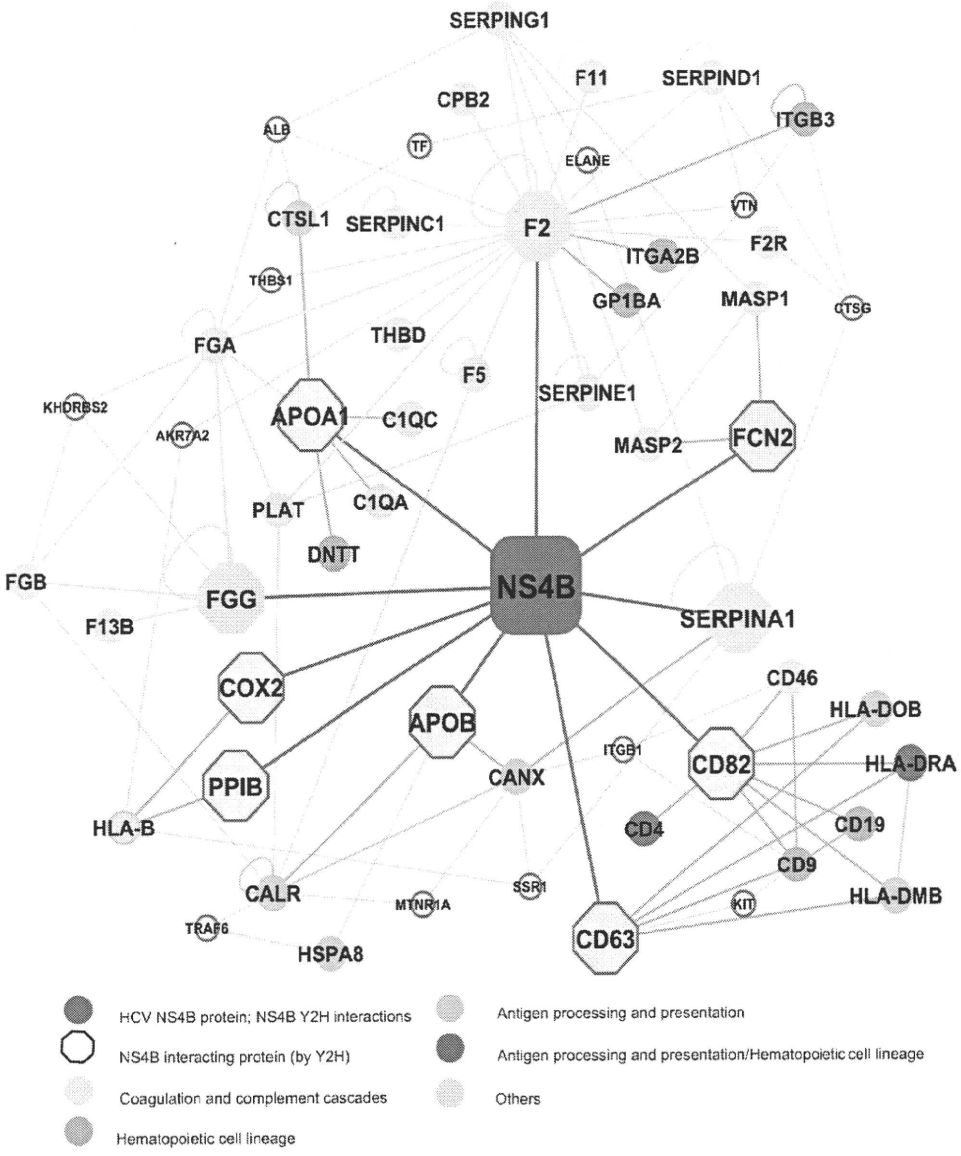

Fig. 5 Network illustration of HCV NS4B interactions associated with KEGG pathways Complement and coagulation cascades, Hematopoietic cell lineage and Antigen processing and presentation. The node sizes differ for better clarity and do not reflect any topological attributes.

immune response and viral persistence. Blocking complement activation can ameliorate the effects of HCV induced hepatic inflammation, suggesting the complement pathway as an attractive target for anti-HCV therapy.[64–67]

Our Y2H screening identified thrombin (F2), serine protease inhibitor SERPINA1 and fibrinogen gamma chain (FGG) as primary interacting partners of NS4B (Table 1, Fig. 2 and 5). Thrombin is implicated in liver cell fibrosis by inducing hepatic stellate cell proliferation,[68] while SERPINA1 defects are implicated in chronic liver disease, hepatitis C and HCC[69–71] and elevated FGG expression and plasma fibrinogen levels are associated with HCC progression.[72] To the best of our knowledge, however, no interactions between these proteins and NS4B had been reported earlier. Interactions with thrombin and SERPINA1 may allow NS4B to directly perturb the host immune response *via* complement activation and thus contribute to HCV pathogenesis and HCC.

2.3.2.2 Hematopoietic development and antigen presentation. Chronic HCV infection induces a reduced natural killer (NK) cell frequency and activity, resulting in impaired cytokine secretion, a stunted immune response and viral persistence.[73] Analysis of the NS4B network revealed an enrichment of proteins (9 of 254, 4%; $p = 0.0213$) mapped to KEGG pathway "Hematopoietic cell lineage", which functions in the generation of NK cells and T and B lymphocytes. The T and B lymphocytes play an important role in innate and adaptive immune response.[74] Our analysis suggested that interactions with host proteins CD63, CD82 and APOA1 (Apolipoprotein A-I) that physically interact with the components of "Hematopoietic cell lineage" may permit NS4B to influence and possibly impair NK cell development and host immune response (Fig. 5).

We also observed an enrichment of proteins (9 of 254, 4%; $p = 0.0155$) mapped to the KEGG pathway "Antigen processing and presentation". Our network analysis showed that these proteins interact with four NS4B interacting proteins: CD63 and CD82 (interacting with HLA-DMB; HLA-DOB and HLA-DRB); APOB (Apolipoprotein B; interacting with CALR, CANX and HSPA8) and APOA1 (interacting with CTSL1 and DNTT) (Fig. 5). The dysfunction of CD63 and CD82 is implicated in cell-to-cell transmission of HIV-I,[75] suggesting that their interaction with NS4B may be crucial to the spread of HCV in the host. APOB associated cholesterol is positively associated with HCV assembly and entry[76,77] thus, NS4B interactions with APOB (and possibly APOA1) may modulate host lipid metabolism and immune response to facilitate HCV pathogenesis and steatosis.[16]

2.3.2.3 Oxidative stress. NS4B overexpression induces ER stress, unfolded protein response (UPR) and production of ROS, which eventually triggers oxidative stress.[78] However, whether NS4B may induce mitochondrial dysfunction by direct associations remains unclear.

Our Y2H screening identified three mitochondrial proteins COX2 (MT-CO2; Cytochrome c oxidase II), ND4 (mitochondrially encoded NADH dehydrogenase 4) and ATP5G2 (ATP synthase subunit C2) to interact with NS4B (Table 1). These proteins map to the KEGG pathway "Oxidative phosphorylation" and are components of the mitochondrial oxidative machinery. Disruptions in ND4 activity[79] and Cytochrome c oxidase deficiency[80] are associated with oxidative stress. Thus, NS4B interactions with the mitochondrial oxidative machinery may allow the virus to influence host oxidative metabolism and potentially induce insulin resistance, steatosis and fibrosis.

The analysis of NS4B interactions may help unravel further associations in HCV infection. Our Y2H screening identified RBP4 (retinol binding protein 4; Table 1) to interact with NS4B, which is suggested to be inversely correlated with chronic HCV infection.[81] However, a precise understanding of the significance of the interactions identified here would be apparent only with further experimental investigations.

2.4 Validation of novel interactions for their role in HCV replication and release

Traditionally, viral and host proteins associated with various steps in HCV lifecycle (internalisation, replication, assembly and release) have been the primary targets in studies focused on anti-HCV strategies. Due to the lack of a suitable model system for HCV infection, cell culture-based systems for HCV RNA replication and infectious viral particle production have been extensively exploited to understand HCV–host interactions and identify potential anti-HCV drug targets.[4,82–84] Our observations by virtue of extended PPI networks suggested novel and potentially crucial roles of host proteins ENO1, PXN, SLC25A5 and VCL (vinculin), an important component of cell–cell junctions,[85] in HCV replication and persistence in the host.

To further explore the roles of these proteins in HCV life cycle, we performed cellular assays to assess the impact of ENO1, PXN, SLC25A5 and VCL siRNA knockdowns on HCV replication and release. Since HCV-production systems established with HCV JFH1 infectious strain (genotype 2a) isolates alone are capable of both efficient replication and production of infectious viral particles,[86,87] JFH1 was used to infect the Huh7OK1 cell line 24 h after transfection with each siRNA (see Materials and Methods). The infected cells were harvested after 72 h post-infection and the expression of each host protein was assessed by qRT-PCR (Fig. 6A). Supernatant viral RNA was significantly decreased by the knockdown of ENO1, but was not affected by the knockdown of PXN in the infected cells, while SLC25A5 knockdown resulted in a slight but statistically significant increase in the amount of the supernatant viral RNA (Fig. 6B). Intracellular viral RNA was significantly reduced by the knockdown of ENO1, but was unaffected by the knockdown of PXN and SLC25A5 (Fig. 6C), suggesting that ENO1 and SLC25A5 regulate HCV replication and assembly/secretion, respectively. VCL knockdown had no significant effect on the intracellular and supernatant virus RNA in the infected cells (data not shown).

To assess the impact of the knockdown of these genes on other HCV genotypes, we repeated the HCV replication assays using Huh-7 cells including HCV replicons derived from JFH1 and Con1 (genotype 1b) infectious strains. Unlike Huh7OK1 cells infected with JFH1, these replicon systems facilitate studies on HCV replication but not infectious virus

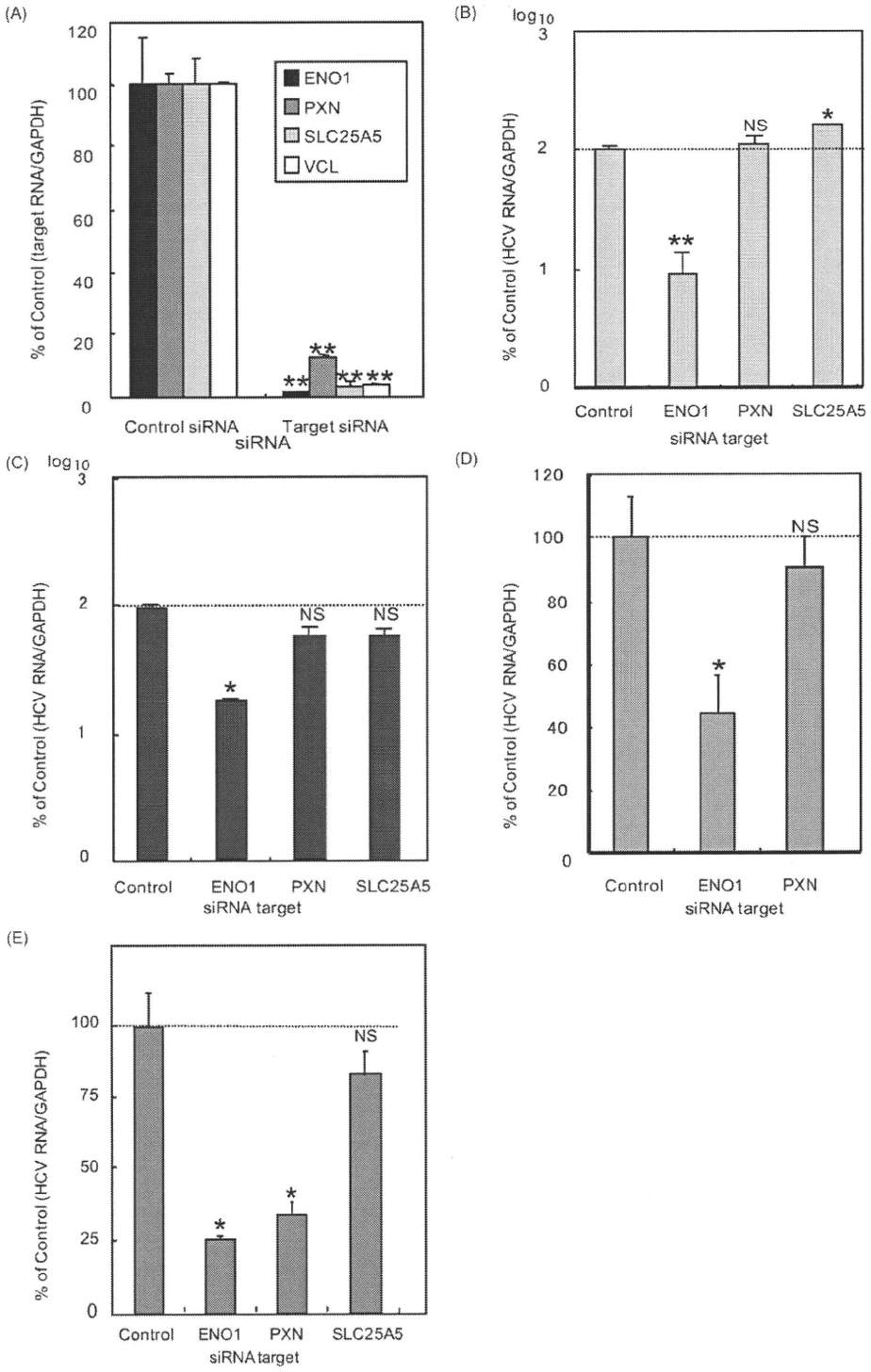

Fig. 6 Effects of knockdown of host protein candidates on HCV propagation and replication. Host proteins ENO1, PXN, SLC25A5 and VCL were suppressed by RNAi in Huh7OK1 cells infected with HCV JFH1 strain (genotype 2a; A, B and C) and in Huh-7 cells including JFH1 replicon RNA (D) or Con1 replicon (genotype 1b; E). Amounts of mRNA of the intracellular host proteins (A), the supernatant viral RNA (B) and the intracellular viral RNA (C, D and E) were estimated. Each value was represented as the percentage of the corresponding quantity measured for the cells transfected with the control siRNA. $*P < 0.05$; $**P < 0.01$; NS: not statistically significant.

secretion.[88] ENO1 knockdown suppressed HCV replication in Huh-7 cells, including either subgenomic replicon RNA of JFH1, which does not encode structural proteins (Fig. 6D), or of Con1, which encodes all viral proteins (Fig. 6E). On the other hand, PXN knockdown in Huh-7 cells including Con1 impaired HCV replication significantly (Fig. 6E), in contrast

to no effect in HuhOK1 cells including JFH1, as described above. This result was not due to the differences in cell lines and the expression of Core protein, because PXN knockdown in Huh-7 cells including JFH1 had no significant effect on intracellular HCV RNA production (Fig. 6D). This observation suggests that PXN possibly regulates the replication of HCV genotype 1b but not of genotype 2a. The standard therapy of pegylated interferon-α plus rebavirin treatment typically achieves less than 50% sustained virological response in HCV genotype 1b infected patients compared to 80% in genotype 2 and 3 infected patients.[89] Therefore, the identification of novel specific (PXN) and non-specific (ENO1) regulators of replication in different HCV genotypes provides potentially attractive targets for developing more effective combinatorial therapies with interferon/RBV treatment.

3. Conclusions

We describe here our observations of PPIs between HCV encoded and host proteins. We first derived a set of experimentally determined interactions between HCV proteins Core and NS4B and host proteins *via* Y2H screens customised for detecting membrane protein interactions. We proceeded to map these interactions onto an overall interaction network that comprised a repertoire of connections potentially required for the two viral proteins to link up with and modulate the components of the host cellular networks. We then employed a network-based approach to further understanding the biological context of these connections in HCV pathogenesis. Core interacting protein SLC25A5 manifested as a potentially important link between the Core protein and the host immune machinery. ENO1, by virtue of gathered interactions, may also function in regulating HCV replication and propagation. We identified 45 previously uncharacterised interactions for NS4B protein that may be crucial for NS4B to function as an important hub in HCV–host interactions. Further investigation of these interactions may help expand substantially our understanding of NS4B function in viral pathogenesis and its potential as an anti-HCV target.

Our observations were then used to prioritise four of the 459 potential candidates in the two extended PPI networks for follow-up experimental investigations through cellular assays based on siRNA knockdowns for HCV genotypes 1b and 2a. These assays validated Core interacting protein ENO1 as a novel regulator of HCV replication and a potentially minor role of SLC25A5 in HCV secretion. In addition, our assays also suggested a genotype 1b specific role of the host protein PXN (which interacts with ENO1 and SLC25A5) in HCV replication. These observations highlighted the attractiveness of the selected host proteins as suitable targets in potentially more effective targeted anti-HCV strategies. The genetic variability of HCV has facilitated the emergence of drug resistance against antiviral drugs that target HCV components. Therefore, antivirals that target less mutable host proteins critical to viral pathogenesis, preferably with minimal adverse side effects, may provide attractive alternatives to existing therapies. That we were able to experimentally validate three of the four genes selected for experimental characterisation reinforces the strengths of elaborate network-based approaches employing diverse functional associations and knowledge-based inputs for identification and prioritisation of suitable targets for experimental and therapeutic investigation. Our study also provides a generic framework for investigating host–pathogen interactions. Such investigation may help identify common themes associated with pathogen (especially viral) infection and help develop effective broad spectrum strategies aimed at ameliorating pathogen (viral) infections.

4. Materials and methods

4.1 Y2H membrane protein assay

Screening for the genes encoding the host proteins that interact with HCV Core and NS4B proteins derived from genotype 1b Con1 strain was performed with a Y2H membrane protein kit system (MoBiTec, Göttingen, Germany) as per the manufacturers specifications. Human adult liver libraries that were constructed based on pPR3 were purchased from MoBiTec and expressed as a fusion protein fused to the N-terminal or the C-terminal end of Nub-G. The cDNA of the Core (or NS4B) encoding region of the HCV polyprotein from the Con1 strain (genotype 1b) was amplified by polymerase chain reaction (PCR) and cloned into the pBT3-N vector (MoBiTec). The screening process was repeated three times to maximise the confidence in the interactions. The total number of screened transformants was 4×10^6, which is about twice the amount of independent clones in the libraries employed for the screening. The clones including genes encoding the Core (or NS4B) interacting proteins were grown on the histidine- and adenosine-deficient culture plate containing a high concentration (10 mM) of 3-amino-1,2,4-triazole (3AT), to remove weak interactions and minimise false positive data. The positive colonies were identified from the blue colour by beta-galactosidase assay (data not shown).

4.2 PPI resources

Secondary interactors of the Core and NS4B interacting proteins were retrieved from BioGRID[90] (version 2.0.63) and PPIView[91] databases. These secondary interactions were merged, filtered for redundancy and appended to the Y2H interactions to infer an extended PPI network. To estimate the robustness of the interactions employed to construct extended PPI networks and infer enriched functional associations (see below), we examined the Human PPI dataset from the I2D database[27] for its overlap with the BioGRID[90]- and PPIView[91]-derived secondary interactions.

4.3 Network topology analysis

Network components were visualised using Cytoscape 2.6,[92] while the network properties such as *node degree distribution* and *average shortest path* measures were computed using Cytoscape NetworkAnalyzer plugin.[93] The degree of node v is defined as the number of nodes directly connected to it, *i.e.*, its first neighbours. Node degree distribution $P(k)$ is the number of nodes with a degree k for $k = 0,1,2\ldots$. By fitting a line on datasets, such as node degree distribution data, the