**Fig. 3** AFP integration value as a predictive factor for HCC. **a** *Flow diagram* showing the number of patients (*n*) classified by baseline alpha-fetoprotein (AFP) levels before interferon (IFN) therapy and average AFP integration value, and the incidence of hepatocellular carcinoma (HCC) of each group. **b** Kaplan–Meier estimates of the incidence of HCC. *Solid line* AFP-low group (AFP levels before interferon therapy <10 ng/mL); *dotted line* AFP high–low group (baseline AFP levels ≥10 ng/mL, average AFP integration value <10 ng/mL); *dashed line* AFP high–high group (both baseline and average AFP integration values ≥10 ng/mL)
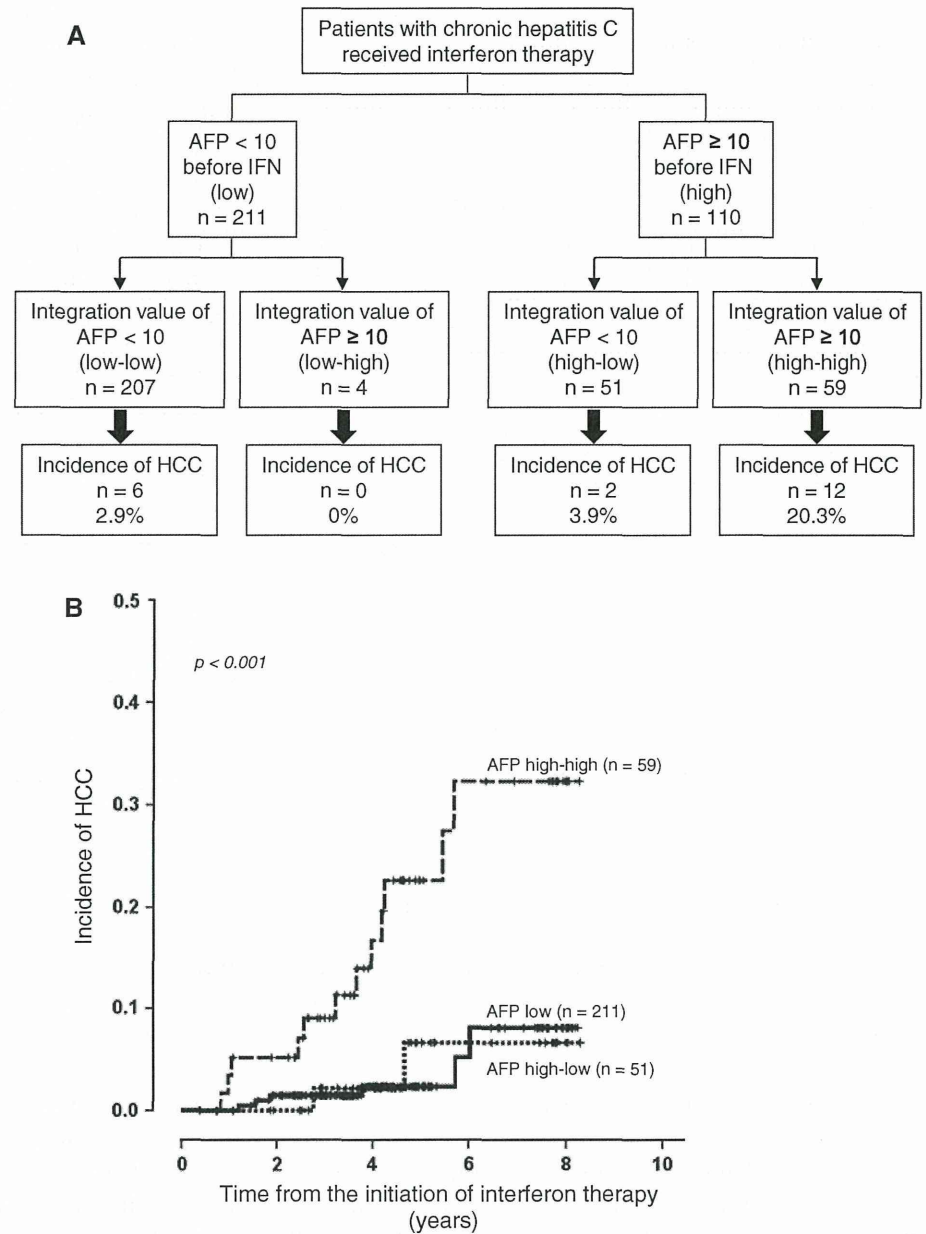
Data from several previous studies suggest that the continuous normalization of alanine aminotransferase (ALT) levels by interferon therapy can reduce the risk of HCC development [36–39]. In addition, one recent study suggested that the ALT integration value is a predictive factor for HCC [35]. In contrast to published data (22), our multivariate analysis did not identify the ALT integration value as a significant predictive factor for HCC incidence, although it was identified as significant by univariate analysis in all 382 patients. Since the previous study did not evaluate AFP levels as a factor for prediction of HCC [35], our results indicate that the AFP integration value is superior to that of ALT as a predictive factor for incidence

of HCC. We do not know the reason for this result, but it is speculated that significance of AFP as a marker of hepatic regeneration resulted in the more accurate prediction of hepatocarcinogenesis by integration value of AFP than that of ALT.

As AFP is a diagnostic marker for the existence of HCC, high integration value of AFP in the present study might be a result of HCC development. However, we concluded that the high AFP integration values in patients who developed HCC were not caused by a result of existence of HCC, because of the following two reasons. First, the last AFP values before detection of HCC were not the highest level in the follow-up periods in 19 of 23 patients who developed

HCC, suggesting that the AFP was not produced by the developing HCC in these patients. Second, to exclude the influence of the remaining four patients whose last AFP levels were the highest in the follow-up periods, we analyzed the same statistical analysis by using average AFP integration values excluded the last two examinations of AFP before the detection of HCC. The results of the analysis also showed average integration value of AFP as a significant predictive factor for incidence of HCC.

Male gender was also identified as an independent risk factor for HCC in non-SVR patients in this study. Several reports have shown that men are at a higher risk of developing HCC than women [6, 10, 33, 40, 41]. The male gender also appears to be a risk factor for more severe disease and a greater risk of developing cirrhosis in chronic hepatitis C [42]. Although the association of male gender with the risk of HCC is as yet unexplained, hormonal or genetic factors may lead to increased risk for HCC and cirrhosis in men as previously discussed [10].

In conclusion, a decrease in the AFP integration value predicts reduced incidence of HCC in patients with hepatitis C receiving interferon therapy. Further prospective studies with a larger number of patients are required to validate the significance of these findings.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Bruix J, Barrera JM, Calvet X, Ercilla G, Costa J, Sanchez-Tapias JM, Ventura M, Vall M, Bruguera M, Bru C, et al. Prevalence of antibodies to hepatitis C virus in Spanish patients with hepatocellular carcinoma and hepatic cirrhosis. Lancet. 1989;2:1004–6.
2. Colombo M, Kuo G, Choo QL, Donato MF, Del Ninno E, Tommasini MA, Dioguardi N, Houghton M. Prevalence of antibodies to hepatitis C virus in Italian patients with hepatocellular carcinoma. Lancet. 1989;2:1006–8.
3. Hasan F, Jeffers LJ, De Medina M, Reddy KR, Parker T, Schiff ER, Houghton M, Choo QL, Kuo G. Hepatitis C-associated hepatocellular carcinoma. Hepatology. 1990;12:589–91.
4. Ikeda K, Saitoh S, Koida I, Arase Y, Tsubota A, Chayama K, Kumada H, Kawanishi M. A multivariate analysis of risk factors for hepatocellular carcinogenesis: a prospective observation of 795 patients with viral and alcoholic cirrhosis. Hepatology. 1993;18:47–53.
5. Tsukuma H, Hiyama T, Tanaka S, Nakao M, Yabuuchi T, Kitamura T, Nakanishi K, Fujimoto I, Inoue A, Yamazaki H, et al. Risk factors for hepatocellular carcinoma among patients with chronic liver disease. N Engl J Med. 1993;328:1797–801.
6. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. Gastroenterology. 2004;127:S35–50.
7. Ikeda K, Marusawa H, Osaki Y, Nakamura T, Kitajima N, Yamashita Y, Kudo M, Sato T, Chiba T. Antibody to hepatitis B core antigen and risk for hepatitis C-related hepatocellular carcinoma: a prospective study. Ann Intern Med. 2007;146:649–56.
8. Liang TJ, Heller T. Pathogenesis of hepatitis C-associated hepatocellular carcinoma. Gastroenterology. 2004;127:S62–71.
9. Yoshida H, Shiratori Y, Moriyama M, Arakawa Y, Ide T, Sata M, Inoue O, Yano M, Tanaka M, Fujiyama S, Nishiguchi S, Kuroki T, Imazeki F, Yokosuka O, Kinoyama S, Yamada G, Omata M. Interferon therapy reduces the risk for hepatocellular carcinoma: national surveillance program of cirrhotic and noncirrhotic patients with chronic hepatitis C in Japan. IHIT Study Group. Inhibition of hepatocarcinogenesis by interferon therapy. Ann Intern Med. 1999;131:174–81.
10. Heathcote EJ. Prevention of hepatitis C virus-related hepatocellular carcinoma. Gastroenterology. 2004;127:S294–302.
11. Lok AS, Seeff LB, Morgan TR, di Bisceglie AM, Sterling RK, Curto TM, Everson GT, Lindsay KL, Lee WM, Bonkovsky HL, Dienstag JL, Ghany MG, Morishima C, Goodman ZD. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. Gastroenterology. 2009; 136:138–48.
12. Effect of interferon-alpha on progression of cirrhosis to hepatocellular carcinoma: a retrospective cohort study. International Interferon-alpha Hepatocellular Carcinoma Study Group. Lancet. 1998;351:1535–9.
13. Camma C, Giunta M, Andreone P, Craxi A. Interferon and prevention of hepatocellular carcinoma in viral cirrhosis: an evidence-based approach. J Hepatol. 2001;34:593–602.
14. Di Bisceglie AM, Shiffman ML, Everson GT, Lindsay KL, Everhart JE, Wright EC, Lee WM, Lok AS, Bonkovsky HL, Morgan TR, Ghany MG, Morishima C, Snow KK, Dienstag JL. Prolonged therapy of advanced chronic hepatitis C with low-dose peginterferon. N Engl J Med. 2008;359:2429–41.
15. Fattovich G, Giustina G, Degos F, Diodati G, Tremolada F, Nevens F, Almasio P, Solinas A, Brouwer JT, Thomas H, Realdi G, Corrocher R, Schalm SW. Effectiveness of interferon alfa on incidence of hepatocellular carcinoma, decompensation in cirrhosis type C. European Concerted Action on Viral Hepatitis (EUROHEP). J Hepatol. 1997;27:201–5.
16. Hayashi K, Kumada T, Nakano S, Takeda I, Kiriyama S, Sone Y, Toyoda H, Shimizu H, Honda T. Incidence of hepatocellular carcinoma in chronic hepatitis C after interferon therapy. Hepatogastroenterology. 2002;49:508–12.
17. Lok AS, Everhart JE, Wright EC, Di Bisceglie AM, Kim HY, Sterling RK, Everson GT, Lindsay KL, Lee WM, Bonkovsky HL, Dienstag JL, Ghany MG, Morishima C, Morgan TR. Maintenance peginterferon therapy and other factors associated with hepatocellular carcinoma in patients with advanced hepatitis C. Gastroenterology. 2011;140:840–9.
18. Nishiguchi S, Kuroki T, Nakatani S, Morimoto H, Takeda T, Nakajima S, Shiomi S, Seki S, Kobayashi K, Otani S. Randomised trial of effects of interferon-alpha on incidence of hepatocellular carcinoma in chronic active hepatitis C with cirrhosis. Lancet. 1995;346:1051–5.
19. Okanoue T, Itoh Y, Minami M, Sakamoto S, Yasui K, Sakamoto M, Nishioji K, Murakami Y, Kashima K. Interferon therapy lowers the rate of progression to hepatocellular carcinoma in chronic hepatitis C but not significantly in an advanced stage: a retrospective study in 1148 patients. Viral Hepatitis Therapy Study Group. J Hepatol. 1999;30:653–9.
20. Izuno K, Fujiyama S, Yamasaki K, Sato M, Sato T. Early detection of hepatocellular carcinoma associated with cirrhosis by combined assay of des-gamma-carboxy prothrombin and alpha-fetoprotein: a prospective study. Hepatogastroenterology. 1995;42:387–93.

21. Trevisani F, D'Intino PE, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P, Domenicali M, De Notariis S, Roda E, Bernardi M. Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. J Hepatol. 2001;34:570–5.

22. Zoli M, Magalotti D, Bianchi G, Gueli C, Marchesini G, Pisi E. Efficacy of a surveillance program for early detection of hepatocellular carcinoma. Cancer. 1996;78:977–85.

23. Alpert E, Feller ER. Alpha-fetoprotein (AFP) in benign liver disease. Evidence that normal liver regeneration does not induce AFP synthesis. Gastroenterology. 1978;74:856–8.

24. Bloomer JR, Waldmann TA, McIntire KR, Klatskin G. Alpha-fetoprotein in noneoplastic hepatic disorders. JAMA. 1975;233: 38–41.

25. Ruoslahti E, Seppala M. Normal and increased alpha-fetoprotein in neoplastic and non-neoplastic liver disease. Lancet. 1972;2: 278–9.

26. Sakurai T, Marusawa H, Satomura S, Nabeshima M, Uemoto S, Tanaka K, Chiba T. *Lens culinaris* agglutinin-A-reactive alpha-fetoprotein as a marker for liver atrophy in fulminant hepatic failure. Hepatol Res. 2003;26:98–105.

27. Taketa K. Alpha-fetoprotein: reevaluation in hepatology. Hepatology. 1990;12:1420–32.

28. Di Bisceglie AM, Sterling RK, Chung RT, Everhart JE, Dienstag JL, Bonkovsky HL, Wright EC, Everson GT, Lindsay KL, Lok AS, Lee WM, Morgan TR, Ghany MG, Gretch DR. Serum alpha-fetoprotein levels in patients with advanced hepatitis C: results from the HALT-C Trial. J Hepatol. 2005;43:434–41.

29. Tateyama M, Yatsuhashi H, Taura N, Motoyoshi Y, Nagaoka S, Yanagi K, Abiru S, Yano K, Komori A, Migita K, Nakamura M, Nagahama H, Sasaki Y, Miyakawa Y, Ishibashi H. Alpha-fetoprotein above normal levels as a risk factor for the development of hepatocellular carcinoma in patients infected with hepatitis C virus. J Gastroenterol. 2011;46:92–100.

30. Murashima S, Tanaka M, Haramaki M, Yutani S, Nakashima Y, Harada K, Ide T, Kumashiro R, Sata M. A decrease in AFP level related to administration of interferon in patients with chronic hepatitis C and a high level of AFP. Dig Dis Sci. 2006;51: 808–12.

31. Tamura Y, Yamagiwa S, Aoki Y, Kurita S, Suda T, Ohkoshi S, Nomoto M, Aoyagi Y. Serum alpha-fetoprotein levels during and after interferon therapy and the development of hepatocellular carcinoma in patients with chronic hepatitis C. Dig Dis Sci. 2009;54:2530–7.

32. Arase Y, Ikeda K, Suzuki F, Suzuki Y, Kobayashi M, Akuta N, Hosaka T, Sezaki H, Yatsuji H, Kawamura Y, Kumada H. Prolonged-interferon therapy reduces hepatocarcinogenesis in aged-patients with chronic hepatitis C. J Med Virol. 2007;79: 1095–102.

33. Asahina Y, Tsuchiya K, Tamaki N, Hirayama I, Tanaka T, Sato M, Yasui Y, Hosokawa T, Ueda K, Kuzuya T, Nakanishi H, Itakura J, Takahashi Y, Kurosaki M, Enomoto N, Izumi N. Effect of aging on risk for hepatocellular carcinoma in chronic hepatitis C virus infection. Hepatology. 2010;52:518–27.

34. Ohno O, Mizokami M, Wu RR, Saleh MG, Ohba K, Orito E, Mukaide M, Williams R, Lau JY. New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. J Clin Microbiol. 1997;35:201–7.

35. Kumada T, Toyoda H, Kiriyama S, Sone Y, Tanikawa M, Hisanaga Y, Kanamori A, Atsumi H, Takagi M, Nakano S, Arakawa T, Fujimori M. Incidence of hepatocellular carcinoma in hepatitis C carriers with normal alanine aminotransferase levels. J Hepatol. 2009;50:729–35.

36. Arase Y, Ikeda K, Suzuki F, Suzuki Y, Kobayashi M, Akuta N, Hosaka T, Sezaki H, Yatsuji H, Kawamura Y, Kumada H. Interferon-induced prolonged biochemical response reduces hepatocarcinogenesis in hepatitis C virus infection. J Med Virol. 2007;79:1485–90.

37. Kasahara A, Hayashi N, Mochizuki K, Takayanagi M, Yoshioka K, Kakumu S, Iijima A, Urushihara A, Kiyosawa K, Okuda M, Hino K, Okita K. Risk factors for hepatocellular carcinoma, its incidence after interferon treatment in patients with chronic hepatitis C. Osaka Liver Disease Study Group. Hepatology. 1998;27:1394–402.

38. Kurokawa M, Hiramatsu N, Oze T, Mochizuki K, Yakushijin T, Kurashige N, Inoue Y, Igura T, Imanaka K, Yamada A, Oshita M, Hagiwara H, Mita E, Ito T, Inui Y, Hijioka T, Yoshihara H, Inoue A, Imai Y, Kato M, Kiso S, Kanto T, Takehara T, Kasahara A, Hayashi N. Effect of interferon alpha-2b plus ribavirin therapy on incidence of hepatocellular carcinoma in patients with chronic hepatitis. Hepatol Res. 2009;39:432–8.

39. Suzuki K, Ohkoshi S, Yano M, Ichida T, Takimoto M, Naitoh A, Mori S, Hata K, Igarashi K, Hara H, Ohta H, Soga K, Watanabe T, Kamimura T, Aoyagi Y. Sustained biochemical remission after interferon treatment may closely be related to the end of treatment biochemical response and associated with a lower incidence of hepatocarcinogenesis. Liver Int. 2003;23:143–7.

40. Kurosaki M, Hosokawa T, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, Tamaki N, Ueda K, Tsuchiya K, Kuzuya T, Nakanishi H, Itakura J, Takahashi Y, Asahina Y, Enomoto N, Izumi N. Hepatic steatosis in chronic hepatitis C is a significant risk factor for developing hepatocellular carcinoma independent of age, sex, obesity, fibrosis stage and response to interferon therapy. Hepatol Res. 2010;40:870–7.

41. Takahashi H, Mizuta T, Eguchi Y, Kawaguchi Y, Kuwashiro T, Oeda S, Isoda H, Oza N, Iwane S, Izumi K, Anzai K, Ozaki I, Fujimoto K. Post-challenge hyperglycemia is a significant risk factor for the development of hepatocellular carcinoma in patients with chronic hepatitis C. J Gastroenterol. 2011;46:790–8.

42. Forns X, Ampurdanes S, Sanchez-Tapias JM, Guilera M, Sans M, Sanchez-Fueyo A, Quinto L, Joya P, Bruguera M, Rodes J. Long-term follow-up of chronic hepatitis C in patients diagnosed at a tertiary-care center. J Hepatol. 2001;35:265–71.

Dynamics of Hepatitis B Virus Quasispecies in Association with Nucleos(t)ide Analogue Treatment Determined by Ultra-Deep Sequencing

Norihiro Nishijima[1], Hiroyuki Marusawa[1]*, Yoshihide Ueda[1], Ken Takahashi[1], Akihiro Nasu[1], Yukio Osaki[2], Tadayuki Kou[3], Shujiro Yazumi[3], Takeshi Fujiwara[4], Soken Tsuchiya[4], Kazuharu Shimizu[4], Shinji Uemoto[5], Tsutomu Chiba[1]

1 Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, 2 Department of Gastroenterology and Hepatology, Osaka Red Cross Hospital, Osaka, Japan, 3 Department of Gastroenterology and Hepatology, Tazuke Kofukai Medical Research Institute, Kitano Hospital, Osaka, Japan, 4 Department of Nanobio Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, Japan, 5 Department of Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan

Abstract

Background and Aims: Although the advent of ultra-deep sequencing technology allows for the analysis of heretofore-undetectable minor viral mutants, a limited amount of information is currently available regarding the clinical implications of hepatitis B virus (HBV) genomic heterogeneity.

Methods: To characterize the HBV genetic heterogeneity in association with anti-viral therapy, we performed ultra-deep sequencing of full-genome HBV in the liver and serum of 19 patients with chronic viral infection, including 14 therapy-naïve and 5 nucleos(t)ide analogue(NA)-treated cases.

Results: Most genomic changes observed in viral variants were single base substitutions and were widely distributed throughout the HBV genome. Four of eight (50%) chronic therapy-naïve HBeAg-negative patients showed a relatively low prevalence of the G1896A pre-core (pre-C) mutant in the liver tissues, suggesting that other mutations were involved in their HBeAg seroconversion. Interestingly, liver tissues in 4 of 5 (80%) of the chronic NA-treated anti-HBe-positive cases had extremely low levels of the G1896A pre-C mutant (0.0%, 0.0%, 0.1%, and 1.1%), suggesting the high sensitivity of the G1896A pre-C mutant to NA. Moreover, various abundances of clones resistant to NA were common in both the liver and serum of treatment-naïve patients, and the proportion of M204VI mutants resistant to lamivudine and entecavir expanded in response to entecavir treatment in the serum of 35.7% (5/14) of patients, suggesting the putative risk of developing drug resistance to NA.

Conclusion: Our findings illustrate the strong advantage of deep sequencing on viral genome as a tool for dissecting the pathophysiology of HBV infection.

Introduction

Hepatitis B virus (HBV) is a non-cytopathic DNA virus that infects approximately 350 million people worldwide and is a main cause of liver-related morbidity and mortality [1–3]. The absence of viral-encoded RNA-dependent DNA polymerase proofreading capacity coupled with the extremely high rate of HBV replication yields the potential to rapidly generate mutations at each nucleotide position within the entire genome [4]. Accordingly, a highly characteristic nature of HBV infection is the remarkable genetic heterogeneity at the inter- and intra- patient level. The latter case of variability as a population of closely-related but nonidentical genomes is referred to as viral quasispecies [5,6]. It is

well recognized that such mutations may have important implications regarding the pathogenesis of viral disease. For example, in chronic infection, G to A point mutation at nucleotide (nt) 1896 in the pre-core (pre-C) region as well as A1762T and G1764A mutations in the core-promoter region are highly associated with HBeAg seroconversion that in general results in the low levels of viremia and consequent clinical cure [7–9]. In contrast, acute infection with the G1896A pre-C mutant represents a high risk for fulminant hepatic failure [10,11]. Although these facts clearly illustrate the clinical implications of certain viral mutation, increasing evidence strongly suggests that

the viral genetic heterogeneity is more complicated than previously thought [12,13].

The major goals of antiviral therapy in patients with HBV infection are to prevent the progression of liver disease and inhibit the development of hepatocellular carcinoma [14]. Oral nucleos(t)ide analogue (NA) have revolutionized the management of HBV infection, and five such antiviral drugs, including lamivudine, adefovir, entecavir, tenofovir, and telbivudine, are currently approved medications [15,16]. These agents are well-tolerated, very effective at suppressing viral replication, and safe, but one of the major problems of NA therapy is that long-term use of these drugs frequently causes the emergence of antiviral drug-resistant HBV due to substitutions at specific sites in the viral genome sequences, which often negates the benefits of therapy and is associated with hepatitis flares and death [16,17]. It is unclear whether viral clones with antiviral resistance emerge after the administration of antiviral therapy or widely preexist among treatment-naïve patients.

There has been a recent advance in DNA sequencing technology [18]. The ultra-deep sequencers allow for massively parallel amplification and detection of sequences of hundreds of thousands of individual molecules. We recently demonstrated the usefulness of ultra-deep sequencing technology to unveil the massive genetic heterogeneity of hepatitis C virus (HCV) in association with treatment response to antiviral therapy [19]. On the other hand, there are a few published studies in which this technology was used to characterize genetic HBV sequence variations [20–22]. Margeridon-Thermet et al reported that the 454 Life Science GS20 sequencing platform provided higher sensitivity for detecting drug-resistant HBV mutations in the serum of patients treated with nucleoside and nucleotide reverse-transcriptase inhibitors [20]. Solmone et al also reported the strong advantage conferred by the same platform to detect minor variants in the serum of patients with chronic HBV infection [21]. Although in these previous studies low-abundant drug-resistant variants were successfully detected, the analyses were focused on the reverse-transcriptase region of circulating HBV in the serum and thus the whole picture of HBV genetic heterogeneity as well as the in vivo dynamics of HBV drug resistant variants in response to anti-viral treatment remains to be clarified. Moreover, intrahepatic viral heterogeneity in patients that achieved the clearance of circulating HBV is largely unknown.

By taking the advantage of an abundance of genetic information obtained by utilizing the Illumina Genome Analyzer II (Illumina, San Diego, CA) as a platform of ultra-deep sequencing, we determined the whole HBV sequence in the liver and serum of patients with chronic HBV infection to evaluate viral quasispecies characteristics. Moreover, we investigated the prevalence of rare drug-resistant HBV variants as well as detailed dynamic changes in the viral genetic heterogeneity in association with NA administration. Based on the abundant genetic information obtained by ultra-deep sequencing, we clarified the precise prevalence of HBV clones with G1896A pre-C mutations in association with HBe serostatus in chronically infected patients with or without NA treatment. We also detected a variety of minor drug-resistant clones in treatment-naïve patients and their dynamic changes in response to entecavir administration, demonstrating the potential clinical significance of naturally-occurring drug-resistant mutations.

Materials and Methods

Ethics Statement

The Kyoto University ethics committee approved the study, and written informed consent for participation in this study was obtained from all patients. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Patients

The liver tissues of 19 Japanese patients that underwent living-donor liver transplantation at Kyoto University due to HBV-related liver disease were available for viral genome analyses. These individuals included 13 men and 6 women, aged 41 to 69 years (median, 55.2 years) and all but one were infected with genotype C viruses. Participants comprised 19 patients with liver cirrhosis caused by chronic HBV infection, including 14 antiviral therapy-naïve cases (chronic-naïve cases) and 5 cases receiving NA treatment, with either lamivudine or entecavir (chronic-NA cases) (Table 1). Serum HBV DNA levels were significantly higher in chronic-naïve cases than in chronic NA cases (median serum HBV DNA levels were 5.6, and <2.6 log copies/ml, respectively, Table 1). Liver tissue samples were obtained at the time of transplantation, frozen immediately, and stored at −80°C until use. Serologic analyses of HBV markers, including hepatitis B surface antigen (HBsAg), antibodies to HBsAg, anti-HBc, HBeAg, and antibodies to HBeAg, were determined by enzyme immuno-assay kits as described previously [23]. HBV DNA in the serum before transplantation was examined using a polymerase chain reaction (PCR) assay (Amplicor HBV Monitor, Roche, Branchburg, NJ). To examine the dynamics of viral quasispecies in response to anti-HBV therapy, paired serum samples of 14 treatment-naïve patients before and after administration of daily entecavir (0.5 mg/day) were subjected to further analyses on viral genome.

Direct population Sanger sequencing

DNA was extracted from the liver tissue and serum using a DNeasy Blood & Tissue Kit (Qiagen, Tokyo, Japan). To define the consensus reference sequences of HBV in each clinical specimen, all samples were first subjected to direct population Sanger sequencing using the Applied Biosystems 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA). Oligonucleotide primers for the HBV genome were designed to specifically amplify whole viral sequences as two overlapping fragments using the sense primer 169_F and antisense primer 2847_R to yield a 2679-bp amplicon (amplicon 1), and the sense primer 685_F and antisense primer 443_R to yield a 2974-bp amplicon (amplicon 2; Table S1). HBV sequences were amplified using Phusion High-Fidelity DNA polymerase (FINZYMES, Espoo, Finland). All amplified PCR products were purified using the QIAquick Gel Extraction kit (Qiagen) after agarose gel electrophoresis and used for direct sequencing. The serum of a healthy HBV DNA-negative volunteer was used as a negative control.

Viral genome sequencing by massively-parallel sequencing

Massively-parallel sequencing with multiplexed tags was performed using the Illumina Genome Analyzer II as described [19]. The end-repair of DNA fragments, addition of adenine to the 3′ ends of DNA fragments, adaptor ligation, and PCR amplification by Illumina PCR primers were performed as described previously [24]. Briefly, the viral genome sequences were amplified by high-fidelity PCR using oligonucleotide primers as described above, sheared by nebulization using 32 psi N2 for 8 min, and then the sheared fragments were purified and concentrated using a QIAquick PCR purification Kit (Qiagen). Nucleotide overhangs resulting from fragmentation were then converted into blunt ends using T4 DNA polymerase and Klenow

Table 1. Characteristics of patients with chronic HBV infection analyzed in this study.

	Chronic-naïve (N = 14)	Chronic-NA (N = 5)
Age[†]	55.5 (41–69)	55.0 (49–68)
Sex (male/female)	9/5	4/1
Alanine aminotransaminase (IU/I)[†]	41 (10–74)	30 (15–65)
Total bilirubin (mg/dl)[†]	0.9 (0.5–31.1)	1.7 (0.6–4.5)
Platelet count (×10⁴/mm³)[†]	12.7 (3.3–27.6)	5.1 (3.6–11.3)
HBV genotype		
B	1	0
C	13	5
Viral load (log copies /ml)[†]	5.6 (<2.6–8.8)*	<2.6 (<2.6–5.3)*
HBe-serostatus (HBeAg+/HBeAb+)	8/6	0/5
Fibrosis		
F0–F2	6	0
F3–F4	8	5
Activity		
A0–A1	7	3
A2–A3	7	2

†Values are median (range).
*P = 0.042.
doi:10.1371/journal.pone.0035052.t001

enzymes, followed by the addition of terminal 3′ A-residues. An adaptor containing unique 6-bp tags, such as "ATCACG" and "CGATGT" (Multiplexing Sample Preparation Oligonucleotide Kit, Illumina), was then ligated to each fragment using DNA ligase. We then performed agarose gel electrophoresis of adaptor-ligated DNAs and excised bands from the gel to produce libraries with insert sizes ranging from 200 to 350 bp. These libraries were amplified independently using a minimal PCR amplification step of 18 cycles by Illumina PCR primers with Phusion High-Fidelity DNA polymerase. The DNA fragments were then purified with a MinElute PCR Purification Kit (Qiagen), followed by quantification using the NanoDrop 2000C (Thermo Fisher Scientific, Waltham, MA) to make a working concentration of 10 nM. Cluster generation and sequencing was performed for 64 cycles on the Illumina Genome Analyzer II according to the manufacturer's instructions. The obtained images were analyzed and base-called using GA pipeline software version 1.4 with the default settings provided by Illumina.

Genome Analyzer sequence data analysis

Using the high performance alignment software "NextGene" (SoftGenetics, State College, PA), the 64 base-pair reads obtained from the Genome Analyzer II were aligned with the reference sequences of 3215 bp that were determined by direct population Sanger sequencing of each clinical specimen. Reads with 90% or more bases matching a particular position of the reference sequences were aligned. Furthermore, two quality filters were used for sequencing reads: the reads with a median quality score of more than 30 and no more than 3 uncalled nucleotides were allowed anywhere in the 64 bases. Only sequences that passed the quality filters, rather than raw sequences, were analyzed and each position of the viral genome was assigned a coverage depth, representing the number of times the nucleotide position was sequenced.

Allele-specific quantitative real-time PCR and semiquantitative PCR to determine the relative proportion of G1896A pre-C mutant

To determine the relative proportion of the G1896A pre-C mutant, allele-specific quantitative real-time PCR was performed based on the previously described method [25,26]. Oligonucleotide primers were designed individually to amplify the pre-C region of wild-type and the G1896A pre-C mutant HBV. Three primers were used for this protocol, two allele-specific sense primers, 1896WT_F (for wild-type) and 1896MT_F (for the G1896A pre-C mutant), and one common antisense primer, 2037_R (Table S1). Quantification of wild-type and the G1896A pre-C mutant was individually performed by real-time PCR using a Light Cycler 480 and Fast Start Universal SYBR Master (Roche, Mannheim, Germany) [27]. The relative proportion of the G1896A pre-C mutant was determined to calculate the G1896A pre-C mutant/total HBV ratios. Performance of this assay was tested using mixtures of two previously described plasmids, pcDNA3-HBV-wt#1 and pcDNA3-HBV-G1896A pre-C mutant [28]. Semiquantitative PCR was performed using primers described above, then agarose gel electrophoresis was performed.

Statistical analysis

Results are expressed as mean or median, and range. Pretreatment values were compared using the Mann-Whitney U-test or the Kruskal Wallis H-test. P values less than 0.05 were considered statistically significant.

The viral quasispecies characteristics were evaluated by analyzing the genetic complexity based on the number of different sequences present in the population. Genetic complexity for each site was determined by calculating the Shannon entropy using the following formula:

$$Sn = -\frac{\sum_{i=1}^{n} fi(lnfi)}{N}$$

where n is the number of different species identified, fi is the observed frequency of a particular variant in the quasispecies, and N is the total number of clones analyzed [12,13]. The mean viral complexity in each sample was determined by calculating the total amounts of the Shannon entropy at each nucleotide position divided by the total nucleotide number (e.g., 3215 bases) of each HBV genome sequence.

Nucleotide sequence accession number

All sequence reads have been deposit in DNA Data Bank of Japan Sequence Read Archive (http://www.ddbj.nig.ac.jp/index-e.html) under accession number DRA000435.

Results

Validation of multiplex ultra-deep sequencing of the HBV genome

To differentiate true mutations from sequencing errors in the determined sequences, we first generated viral sequence data from the expression plasmid, pcDNA3-HBV-wt#1, encoding wild-type genotype C HBV genome sequences [28]. For this purpose, we determined the PCR-amplified HBV sequences derived from the expression plasmid using high-fidelity Taq polymerase to take the PCR-induced errors as well as sequencing errors into consideration. Viral sequences determined by the conventional Sanger method were used as reference sequences for aligning the amplicons obtained by ultra-deep sequencing. Three repeated ultra-deep sequencing generated a mean of 77,663 filtered reads, corresponding to a mean coverage of 38,234 fold at each nucleotide site (Table S2). Errors comprised insertions (0.00003%), deletions (0.00135%), and nucleotide mismatches (0.037%). The mean overall error rate was 0.034% (distribution of per-nucleotide error rate ranged from 0 to 0.13%) for the three control experiments, reflecting the error introduced by high-fidelity PCR amplification and by multiplex ultra-deep sequencing that remained after filtering out problematic sequences. We also confirmed that multiplex ultra-deep sequencing with and without the high-fidelity PCR amplification with HBV-specific primer sets showed no significant differences in the error rates on the viral sequence data (mean error rate 0.034% vs 0.043%). Accordingly, we defined the cut-off value in its current platform as 0.3%, a value nearly 1 log above the mean overall error rate.

Next, we performed additional control experiments to verify the detectability of the low abundant mutations that presented at a frequency of less than 0.3%. For this purpose, we introduced expression plasmids with a single-point mutation within that encoding a wild-type viral sequence with a ratio of 1:1000 and assessed the sensitivity and accuracy of quantification using high-fidelity PCR amplification followed by multiplex ultra-deep sequencing in association with the different coverage numbers (Table S3). Repeated control experiments revealed that the threshold for detecting low-abundant mutations at an input ratio of 0.10% among the wild-type sequences ranged between 0.11% and 0.24%, indicating that there was no significant difference in the detection rate or error rates under the different coverage conditions. Based on these results, the accuracy of ultra-deep sequencing in its current platform for detecting low-level viral mutations was considered to be greater than 0.30%.

Viral complexity of the HBV quasispecies in association with clinical status

To clarify HBV quasispecies in association with clinical status, we performed multiplex ultra-deep sequencing and determined the HBV full-genome sequences in the liver and serum with

chronic HBV infection. First, we compared the sequences of the viral genome determined in the liver tissue with those in the serum and found no significant differences in the viral population between the liver and serum of the same individual. Indeed, the pattern and distribution of genetic heterogeneity of the viral nucleotide sequences in the liver tissue were similar to those observed in the serum of the same patient (Figure S1), suggesting that a similar pattern of viral heterogeneity was maintained in the liver and serum of patients with chronic HBV infection.

Next, we compared the viral heterogeneity in the liver of chronic-naïve and chronic-NA cases. A mean of 5,962,996 bp nucleotides in chronic-naïve cases and 4,866,783 bp nucleotides in chronic-NA cases were mapped onto the reference sequences, and an overall average coverage depth of 1,855 and 1.514 was achieved for each nucleotide site of the HBV sequences, respectively (Table 2). The frequencies of mutated positions and altered sequence variations detected in each viral genomic region are summarized in Table 2. The overall mutation frequency of the total viral genomic sequences was determined to be 0.87% in chronic-naïve cases and 0.69% in chronic-NA cases. Most genomic changes observed in viral variants were single base substitutions, and the genetic heterogeneity of the viral nucleotide sequences was equally observed throughout the individual viral genetic regions, including the pre-surface (preS), S, pre-core~core (preC-C), and X (Table 2). Consistent with the findings obtained from the viral mutation analyses, the overall viral complexity determined by the Shannon entropy value was 0.047 in chronic-naïve and 0.036 in chronic-NA cases, and the viral complexity was equally observed throughout the individual viral genetic region (Figure 1A). Among chronic-naïve cases, we observed no significant differences in the viral complexity in HBV DNA level, age, or degree of fibrosis (Figure 1B).

High sensitivity of the G1896A pre-C mutant to nucleos(t)ide analogues

Emergence of G1896A mutation in the pre-C region, and A1762T and G1764A mutations in the core-promoter region is well known to be associated with HBe-seroconversion [7–9]. We then evaluated the prevalence of these three mutations in the chronically HBV-infected liver, in association with HBe serologic status and the NA treatment history. In chronic-naive cases, 6 and 8 patients showed the pre- and post- HBeAg seroconversion status, respectively (Table 3). The mean prevalence of the G1896A pre-C mutant in HBeAg-positive cases was lower than that in the HBeAg-negative cases (27.4% and 46.5%, respectively). Importantly, however, 4 of 8 HBeAg-negative cases showed a relatively low prevalence of the G1896A pre-C mutant (Liver #8, #12, #13, #14), and all but one case (Liver #10) showed a high prevalence of the A1762T and G1764A mutations, irrespective of HBe serologic status and NA treatment history (Table 3). These findings suggested that other mutations except G1896A, A1762T and G1764A were also involved in the HBeAg seroconversion status. Notably, liver tissues of all but one (Liver #17) chronic-NA cases showed extremely low levels of the G1896A pre-C mutant (0.0, 0.0, 0.1, and 1.1%), suggesting the high sensitivity of the G1896A pre-C mutant to NA (Table 3).

To confirm the difference of the sensitivity to NA between the wild-type and the G1896A pre-C mutant, we examined the dynamic changes of the relative proportion of the G1896A pre-C mutant in the serum of 14 treatment-naïve patients before and after entecavir administration. Consistent with the findings obtained by ultra-deep sequencing, quantitative real-time PCR revealed that entecavir administration significantly reduced the proportion of the G1896A pre-C mutant in 13 of 14 cases (92.9%)
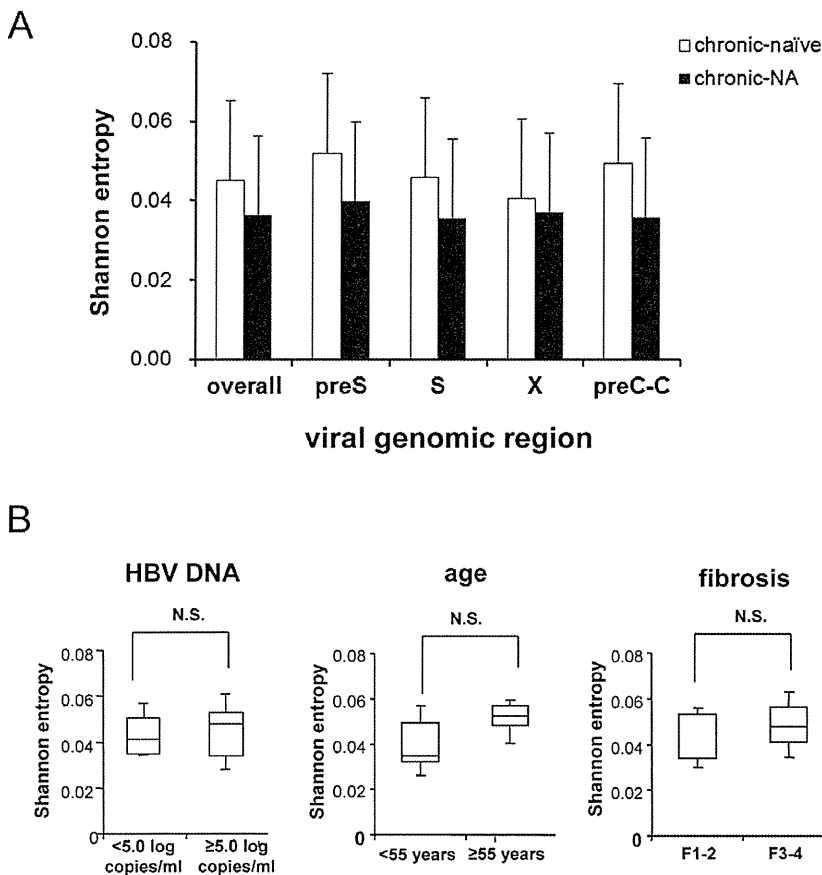
Figure 1. Viral complexity of the HBV quasispecies in association with clinical status. (A) The Shannon entropy values for each viral genomic region were determined in the liver of chronic-naïve and chronic-NA cases. (B) Among the chronic-naïve cases, the Shannon entropy values are shown for patients with serum HBV DNA levels less than 5.0 log copies/ml (<5.0) and greater than 5.0 log copies/ml (≥5.0) (left panel), patients under the age of 55 years (<55) and over the age of 55 (≥55) (middle panel), and patients with low (F1-2) and high (F3-4) liver fibrosis levels (right panel). preS: pre-surface, preC-C: precore~core N.S.: not significant.
doi:10.1371/journal.pone.0035052.g001

Table 2. The frequency of mutation rate and the Shannon entropy in each viral genome region.

	Liver	
	Chronic-naïve (N = 14)	Chronic-NA (N = 5)
Average aligned reads	93,172	76,043
Average aligned nucleotides	5,962,996	4,866,783
Average coverage	1,855	1,514
Mutation rate (%)		
Overall	0.87	0.69
preS	0.92	0.81
S	0.96	0.71
preC-C	1.05	0.72
X	0.63	0.61
Shannon entropy	0.047	0.036

Mutation rate (%): the ratio of total different nucleotides from the reference sequence to total aligned nucleotides.
preS: pre-surface, preC-C: pre-core~core.
doi:10.1371/journal.pone.0035052.t002

irrespective of their HBeAg serostatus, while the G1896A pre-C mutant were detectable in substantial proportion before treatment in all cases (Figure 2A, 2B and 2C; p = 0.001). These results further support the findings that HBV clones comprising the G1896A mutation were more sensitive to NA than those with wild-type sequences.

Prevalence of drug-resistant HBV clones in the liver of treatment-naïve patients

Increasing evidence suggests that drug-resistant viral mutants can be detected in the serum of treatment-naïve patients with chronic HBV infection [20,21]. Thus, we next determined the actual prevalence of spontaneously-developed drug-resistant mutants in chronically-infected liver of treatment-naïve patients to evaluate whether NA treatment potentiates the expansion of drug-resistant clones. The drug-resistant mutations examined included two mutations resistant to lamivudine and entecavir, four mutations resistant to entecavir, and three mutations resistant to adefovir [16,17]. Based on the detection rate of the low-level viral clones determined by the control experiments, we identified the drug-resistant mutants present in each specimen at a frequency of more than 0.3% among the total viral clones. Based on these criteria, at least one resistant mutation was detected in the liver of all of the chronic-naïve cases with chronic HBV infection (Table 4).

Table 3. The prevalence of G1896A mutation in the pre-C region, and A1762T and G1764A mutations in the core-promoter region in the liver of patients chronically infected with HBV.

	HBeAg/HBeAb	NA (duration of treatment)	Mutation Frequency					
			G1896A (Pre C)		A1762T (CP)		G1764A (CP)	
Chronic-naive								
Liver #1	+/−	-	640/1652	(38.7)	1647/1941	(84.9)	1683/1979	(85.0)
Liver #2	+/−	-	9/596	(1.5)	682/687	(99.3)	683/689	(99.1)
Liver #3	+/−	-	273/672	(40.6)	767/769	(99.7)	757/760	(99.6)
Liver #4	+/−	-	204/701	(29.1)	610/625	(97.6)	602/621	(96.9)
Liver #5	+/−	-	27/152	(17.8)	249/250	(99.6)	245/248	(98.8)
Liver #6	+/−	-	228/621	(36.7)	727/729	(99.7)	743/744	(99.9)
Liver #7	−/+	-	740/1193	(62.0)	1908/1913	(99.7)	1888/1913	(98.7)
Liver #8	−/+	-	111/1892	(5.9)	2321/2325	(99.8)	2335/2339	(99.8)
Liver #9	−/+	-	10935/10944	(99.9)	12019/12032	(99.9)	12163/12170	(99.9)
Liver #10	−/+	-	4554/4593	(99.2)	1/5191	(0)	4/5188	(0.1)
Liver #11	−/+	-	811/921	(88.1)	1234/1236	(99.8)	1226/1228	(99.8)
Liver #12	−/+	-	93/1265	(7.4)	1234/1234	(100)	1228/1229	(99.9)
Liver #13	−/+	-	83/877	(9.5)	1465/1529	(95.8)	1485/1549	(95.9)
Liver #14	−/+	-	0/717	(0)	1078/1410	(76.5)	1089/1414	(77.0)
Chronic-NA								
Liver #15	−/+	LAM (156w)	0/390	(0)	441/453	(97.4)	435/448	(97.1)
Liver #16	−/+	ETV (1w)	0/1399	(0)	1624/1632	(99.5)	1625/1630	(99.7)
Liver #17	−/+	LAM (144w)	345/816	(42.3)	988/991	(99.7)	994/994	(100)
Liver #18	−/+	LAM (98w)	2/3963	(0.1)	1015/1188	(85.4)	1190/1194	(99.7)
Liver #19	−/+	LAM (11w)	48/4214	(1.1)	3438/3456	(99.5)	3446/3462	(99.5)

Values in parenthesis show mutation frequency (%): the ratio of total mutant clones to total aligned coverage at each nucleotide sites.
NA: nucleotide analogue, pre C: precore, CP: core promoter, LAM: lamivudine, ETV: entecavir.
doi:10.1371/journal.pone.0035052.t003

The prevalence of the 9 drug-resistant mutations detected by ultra-deep sequencing in 14 chronic-naïve cases ranged from 0.3% to 30.0%, indicating that the proportion of resistant mutations substantially differed in each case. The most commonly detected mutation was M204VI (9 cases) and M250VI (11 cases), which were resistant to lamivudine and entecavir, and entecavir, respectively. Other mutations resistant to adefovir were detected in 7 (50.0%) and 3 (21.4%) cases at A181TV and N236T, respectively (Table 4).

Nine (64.2%) chronic-naïve cases possessed the M204VI mutants in their liver tissues and the proportion of mutant clones among the totally infected viruses ranged from 0.3% to 1.1% among the M204VI mutant-positive patients. In chronic-NA cases, 4 of 5 (80.0%) liver tissues harbored the M204VI mutants with the proportion among the totally infected viruses ranging from 0.4% to 18.7% (Table 4), while the mean serum HBV DNA was suppressed below 2.6 log copies/ml (Table 1). These results suggest that the mutant HBV clones comprising various drug-resistant mutations could latently exist even in the liver of NA treatment-naïve cases.

Expansion of drug-resistant HBV clones harboring M240VI mutations in response to NA administration

To clarify the risk of latent expansion of drug-resistant mutations due to NA treatment, we next examined the early dynamic changes of the prevalence of M204VI mutants in the serum of treatment-naïve patients in response to entecavir treatment. Ultra-deep sequencing provided a mean 40,791- and 38,823-fold coverage of readings, which were mapped to the M204VI nucleotide position at the YMDD sites of each reference sequence in patients before and after entecavir treatment.

Five of 14 (35.7%) patients harbored the M204VI mutations prior to entecavir treatment. Although the serum HBV DNA levels were significantly reduced in response to entecavir in all cases, the M204VI mutant clones were detected in 9 cases (64.3%) after entecavir administration (Table 5). Notably, one patient (Serum #3) who harbored the M240VI mutant clones at baseline had a relatively large expansion of drug-resistant clones among the total viral population in a time-dependent manner in response to entecavir treatment (Table 5). Similarly, M240VI mutant clones became detectable after entecavir administration in four patients (Serum #1, #7, #12, #13) that harbored no resistant mutants at baseline (Table 5). We found no correlation between the degree of the increase in the relative prevalence of M204VI mutant clones and that of the reduction in serum HBV DNA levels. Although only a limited number of patients exhibited a substantial increase in M204VI mutant clones after administration of anti-viral therapy, our findings might suggest that entecavir treatment latently causes selective survival of drug-resistant mutants in treatment naïve patients with chronic HBV infection.
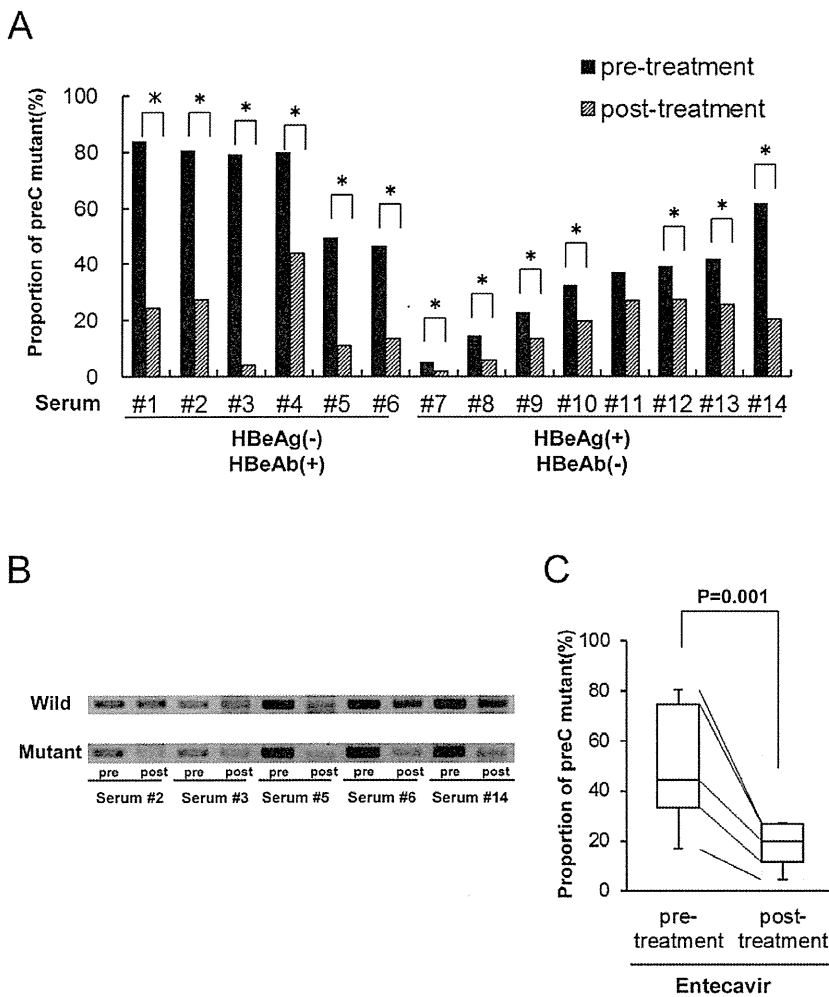
A

B

C

Figure 2. The reduction in the relative proportion of the G1896A pre-C mutant clones after entecavir administration. (A) The relative proportion of the G1896A pre-C mutant was determined in the serum of treatment-naïve patients pre- and post-entecavir administration using quantitative real-time PCR. Serum #1~6 were HBeAg-negative and HBeAb-positive, and Serum #7~14 were HBeAg-positive and HBeAb-negative before treatment. *: p<0.05 (B) Semiquantitative PCR analysis was performed using primers specific to the wild-type (upper panel) or G1896A pre-C mutant (lower panel) pre- and post-entecavir administration. A representative result from 5 cases is shown. (C) The relative proportion of the G1896A pre-C mutant was compared in 14 treatment-naïve patients between pre- and post-entecavir administration.
doi:10.1371/journal.pone.0035052.g002

Discussion

Direct population sequencing is the most common method for detecting viral mutations [29]. Conventional sequencing techniques, however, are not efficient for evaluating large amounts of genetic information of the viruses. Newly developed ultra-deep sequencing technology have revolutionized genomic analyses, allowing for studies of the dynamics of viral quasispecies as well as rare genetic variants of the viruses that cannot be detected using standard direct population sequencing techniques [30,31]. The sensitivity of ultra-deep sequencing analysis is primarily limited by errors introduced during PCR amplification and the sequencing reaction, thus it is a challenge to distinguish rare variants from sequencing artifacts. In the present study, we optimized the ultra-deep sequencing with a multiplex-tagging method and reproducibly detected variants within HBV quasispecies that were as rare as 0.3%. Based on this ultra-deep sequencing platform, we determined the abundant genetic heterogeneity of HBV at the intra- and inter-individual levels.

Because of its ability to handle abundant viral genome information, ultra-deep sequencing allowed us to evaluate low-abundant virus variants of patients with chronic HBV infection in detail. It is widely accepted that HBe seroconversion is highly associated with the emergence of G1896A pre-C and/or A1762T and G1764A core promoter mutant clones [7–9]. Unexpectedly, however, our results showed a diverse range of G1896A frequency (0–99.9%) in HBeAg-negative subjects and a high prevalence of core promoter mutations, irrespective of HBe serostatus. Consistent with our observation, previous studies utilizing conventional sequencing methods reported that the frequency of the G1896A pre-C mutant ranged from 12% to 85% [32]. All but one patient (Liver #10) showing a predominance of A1762T and G1764A were infected with genotype C, while patient#10 was infected with genotype B. Because A1762T and G1764A are reported to be significantly more frequent in genotype C [33], the difference in the prevalence of A1762T and G1764A in our study might be a reflection of the viral HBV genotype rather than HBe serostatus. Further investigation of the actual prevalence of these mutations