probe sets. RNA samples from the cerebrum, cerebellum, liver, and testis of a 3-year-old male cynomolgus monkey were extracted using TRIZOL (Invitrogen) and hybridized to the GeneChip with duplication in a single experiment. The *M. fascicularis* GeneChip contains at most 11 perfect-match probes (25-mers complete matches to the cDNA sequences) and 11 mismatch probes (containing one mismatched oligo) for each probe set, similar to other GeneChip formats. Normalization, signal detection, and signal intensity calculation of the microarrays were performed using Affymetrix MAS5.0 software. Transcripts were considered as expressed when the probe set of both the duplicates agreed for the significant expression ($P \leq 0.05$) [33]. The raw array data were deposited in Gene Expression Omnibus [GEO: GSM201873–201880]. The array design and the sequences of oligonucleotide probes were deposited in the public database [GEO: GPL5396].

RT-PCR

Templates of the human brain RNA were purchased from Clontech. The macaque brain RNA was obtained from a 21-year-old male cynomolgus monkey. One microgram of total mRNA was amplified using the PrimeSTAR® RT-PCR Kit (TakaraBio). The temperature and time schedules were 30 cycles at 94°C for 20 s, 60°C for 30 s, and 72°C for 1 min. All primer sequences are presented in Additional file 6.

Human-cynomolgus cDNA sequence alignment

Human-macaque orthologous gene pairs were assigned by the reciprocal best BLAST hit with an e-60 cut-off value. We aligned only that part of the coding sequences (CDS) that was homologous to a BLAST search, because representative human and macaque cDNAs do not necessarily have the same splicing isoforms. The sequences of human and macaque cDNAs were aligned using CLUSTAL W [51], and an unaligned macaque nucleotide was marked by the letter X in the database. Alignments shorter than 100 bp (≤ 33 codons) were filtered for further analysis. In the database, the positions including deletion in the human sequence (or insertion in the macaque sequence) were dropped for estimating the substitution rates. The non-synonymous substitution rate per non-synonymous site (*Ka*) and the synonymous substitution rate per synonymous site (*Ks*) were estimated using the Li-Pamilo-Bianchi method [52,53]. *Ka*/*Ks* ratios were set to 100 in the database when the *Ks* value was zero.

Human-rhesus-cynomolgus cDNA sequence alignment

The predicted cDNA sequences of rhesus macaques were downloaded from Ensembl (MMUL1.0) and aligned with the human RefSeq sequences. Orthology between the rhesus and cynomolgus genes was confirmed again using the cynomolgus-rhesus reciprocal BLAST hit, and human-rhesus-cynomolgus cDNA alignments were compiled. Alignments containing any frameshifting indels and those shorter than 100 bp (≤ 33 codons) were filtered, which resulted in 2655 alignments (dataset I). The rhesus cDNA sequences were then mapped on the draft genome sequence of the rhesus macaque (rheMac2). The rhesus macaque genes showing > 80% homology to more than one locus on the rhesus genome were removed from the alignments, which yielded 1499 human-rhesus-cynomolgus cDNA alignments (dataset II). To estimate the divergence among the three species, *Ka*(*dn*) and *Ks*(*ds*) were estimated using the maximum likelihood method implemented in the PAML program package [54]. We estimated the transition/transversion ratios in 4-fold degenerated sites using the concatenated cynomolgus and rhesus alignments in advance, and fixed the value to the observed value. The test of positive selection was conducted using the branch-site test of positive selection described by Zhang et al. [41], applying the critical values of 2.71 and 5.41 at 5% and 1% significance level without a Bonferroni correction, respectively.

Estimation of the divergence time between cynomolgus and rhesus macaques

Maximum likelihood estimation (MLE) of the divergence time and ancestral population size was performed using the method of Takahata and Satta [11,14]. MLE was determined using the Newton-Raphson algorithm with many possible initial values. The standard error was determined from the numerically evaluated Fisher information matrix. In order to correct a PCR error rate, we estimated the PCR error rate to be 5.40 × 10-4, which was derived from the difference in the synonymous substitution rates of the cynomolgus and rhesus lineages. We assumed that the generation time and the effect of selection on the synonymous sites of the two macaques were the same, and that the erroneous nucleotide incorporated by PCR did not skew. Therefore, when a synonymous substitution in the cynomolgus lineage was found, it was considered that the substitution is because of the PCR error with a probability of 0.178 (5.40 × 10-4/3.04 × 10-3). We randomly corrected the number of substitutions in the raw data, generated pseudo data for 1000 times, and estimated the evolutionary parameter for each time.

Abbreviations

CDS, coding sequence; EST, expressed sequence tag; MLE, maximum likelihood estimation; ORF, open reading frame; UTR, untranslated region; MRCA, most recent common ancestor

Authors' contributions

NO contributed to the designing of the research, performed the experiments and data analysis, and wrote the manuscript. KH, KT, and JK designed the research and contributed to the manuscript. M Hirata performed the

computational analysis. YK contributed to the microarray experiments. RT, YU, H, and IT were involved in the cDNA sequencing. M Hida, YS and SS constructed the oligo-capped cDNA libraries. All authors read and approved the final manuscript.

Additional material

Additional file 1

Expression of novel macaque transcripts. Significance of gene expression in the M. fascicularis *oligonucleotide microarray analysis is shown.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-90-S1.xls]

Additional file 2

Unidentified UTR regions in the macaque cDNAs. The regions of macaque cDNAs that did not show homology to human cDNAs are listed.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-90-S2.xls]

Additional file 3

Divergence among the human, cynomolgus, and rhesus genes (dataset 1: without a duplication filtering). Estimation of gene divergence using 2655 human-rhesus-cynomolgus alignment is shown.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-90-S3.doc]

Additional file 4

LRT (likelihood ratio test) statistics for the test of positive selection. Candidate genes under positive selection using branch-site test of positive selection are given with log-likelihood ratio.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-90-S4.xls]

Additional file 5

Specific probes for known genes in the M. fascicularis *microarray. These genes are not found in the previously published macaque oligonucleotide microarray.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-90-S5.xls]

Additional file 6

Primer sequences for RT-PCR. The list shows the primer sequences that were used for RT-PCR.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-90-S6.xls]

Acknowledgements

This study was supported by a Health Science Research Grant from the Ministry of Health, Labor and Welfare of Japan and a Grant for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (19770073).

References

1. Carlsson HE, Schapiro SJ, Farah I, Hau J: **Use of primates in research: a global overview.** *Am J Primatol* 2004, 63:225-237.
2. Melnick DJ, Hoelzer GA, Absher R, Ashley MV: **mtDNA diversity in rhesus monkeys reveals overestimates of divergence time and paraphyly with neighboring species.** *Mol Biol Evol* 1993, 10:282-295.
3. Hayasaka K, Fujii K, Horai S: **Molecular phylogeny of macaques: implications of nucleotide sequences from an 896-base pair region of mitochondrial DNA.** *Mol Biol Evol* 1996, 13:1044-1053.
4. Magness CL, Fellin PC, Thomas MJ, Korth MJ, Agy MB, Proll SC, Fitzgibbon M, Scherer CA, Miner DG, Katze MG, Iadonato SP: **Analysis of the Macaca mulatta transcriptome and the sequence divergence between Macaca and human.** *Genome Biol* 2005, 6:R60.
5. Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, Siepel A, Sikela JM, Attaway T, Bell S, Bernard KE, Buhay CJ, Chandrabose MN, Dao M, Davis C, Delehaunty KD, Ding Y, *et al.*: **Evolutionary and biomedical insights from the rhesus macaque genome.** *Science* 2007, 316:222-234.
6. Dutrillaux B, Biemont MC, Viegas Pequignot E, Laurent C: **Comparison of the karyotypes of four Cercopithecoidae: Papio papio, P. anubis, Macaca mulatta, and M. fascicularis.** *Cytogenet Cell Genet* 1979, 23:77-83.
7. Tosi AJ, Morales JC, Melnick DJ: **Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys.** *Evolution Int J Org Evolution* 2003, 57:1419-1435.
8. Ferguson B, Street SL, Wright H, Pearson C, Jia Y, Thompson SL, Allibone P, Dubay CJ, Spindel E, Norgren RB Jr: **Single nucleotide polymorphisms (SNPs) distinguish Indian-origin and Chinese-origin rhesus macaques (*Macaca mulatta*).** *BMC Genomics* 2007, 8:43.
9. Hernandez RD, Hubisz MJ, Wheeler DA, Smith DG, Ferguson B, Rogers J, Nazareth L, Indap A, Bourquin T, McPherson J, Muzny D, Gibbs R, Nielsen R, Bustamante CD: **Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques.** *Science* 2007, 316:240-243.
10. Smith DG, McDonough JW, George DA: **Mitochondrial DNA variation within and among regional populations of longtail macaques (*Macaca fascicularis*) in relation to other species of the fascicularis group of macaques.** *Am J Primatol* 2007, 69:182-198.
11. Edwards SV, Beerli P: **Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies.** *Evolution Int J Org Evolution* 2000, 54:1839-1854.
12. Takahata N, Satta Y: **Evolution of the primate lineage leading to modern humans: phylogenetic and demographic inferences from DNA sequences.** *Proc Natl Acad Sci USA* 1997, 94:4811-4815.
13. Tosi AJ, Morales JC, Melnick DJ: **Y-Chromosome and Mitochondrial Markers in Macaca fascicularis Indicate Introgression with Indochinese M. mulatta and a Biogeographic Barrier in the Isthmus of Kra.** *Int J Primatol* 2002, 23:161-178.
14. Osada N, Wu CI: **Inferring the mode of speciation from genomic data: a study of the great apes.** *Genetics* 2005, 169:259-264.
15. Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R, Wakamatsu A, Hayashi K, Sato H, Nagai K, Kimura K, Makita H, Sekine M, Obayashi M, Nishi T, Shibahara T, Tanaka T, Ishii S, Yamamoto J, Saito K, Kawai Y, Isono Y, Nakamura Y, Nagahari K, Murakami K, Yasuda T, Iwayanagi T, Wagatsuma M, Shiratori A, Sudo H, *et al.*: **Complete sequencing and characterization of 21,243 full-length human cDNAs.** *Nat Genet* 2004, 36:40-45.
16. Osada N, Hida M, Kusuda J, Tanuma R, Iseki K, Hirata M, Suto Y, Hirai M, Terao K, Suzuki Y, *et al.*: **Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes.** *Gene* 2001, 275:31-37.
17. Osada N, Hida M, Kusuda J, Tanuma R, Hirata M, Hirai M, Terao K, Suzuki Y, Sugano S, Hashimoto K: **Prediction of unidentified human genes on the basis of sequence similarity to novel**

cDNAs from cynomolgus monkey brain. *Genome Biol* 2002, 3:RESEARCH0006.
18. Osada N, Hida M, Kusuda J, Tanuma R, Hirata M, Suto Y, Hirai M, Terao K, Sugano S, Hashimoto K: **Cynomolgus monkey testicular cDNAs for discovery of novel human genes in the human genome sequence.** *BMC Genomics* 2002, 3:36.
19. Osada N, Kusuda J, Hirata M, Tanuma R, Hida M, Sugano S, Hirai M, Hashimoto K: **Search for genes positively selected during primate evolution by 5'-end-sequence screening of cynomolgus monkey cDNAs.** *Genomics* 2002, 79:657-662.
20. Osada N, Hirata M, Tanuma R, Kusuda J, Hida M, Suzuki Y, Sugano S, Gojobori T, Shen CK, Wu CI, Hashimoto K: **Substitution rate and structural divergence of 5'-UTR evolution: comparative analysis between human and cynomolgus monkey cDNAs.** *Mol Biol Evol* 2005, 22:1976-1982.
21. Wang HY, Chien HC, Osada N, Hashimoto K, Sugano S, Gojobori T, Chou CK, Tsai SF, Wu CI, Shen CK: **Rate of Evolution in Brain-Expressed Genes in Humans and Other Primates.** *PLoS Biol* 2007, 5:e13.
22. Maruyama K, Sugano S: **Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides.** *Gene* 1994, 138:171-174.
23. Maglott D, Ostell J, Pruitt KD, Tatusova T: **Entrez Gene: gene-centered information at NCBI.** *Nucleic Acids Res* 2005, 33:D54-58.
24. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, 25:25-29.
25. **QFbase** [http://genebank.nibio.go.jp/qfbase/]
26. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL: **GenBank.** *Nucleic Acids Res* 2007, 35:D21-25.
27. Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T, Down T, Dyer SC, Fitzgerald S, Fernandez-Banet J, Graf S, Haider S, Hammond M, Herrero J, Holland R, Howe K, Howe K, Johnson N, Kahari A, Keefe D, Kokocinski F, Kulesha E, Lawson D, Longden I, Melsopp C, Megy K, *et al.*: **Ensembl 2007.** *Nucleic Acids Res* 2007, 35:D610-617.
28. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA: **Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders.** *Nucleic Acids Res* 2005, 33:D514-517.
29. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, Yura K, Miyazaki S, Ikeo K, Homma K, Kasprzyk A, Nishikawa T, Hirakawa M, Thierry-Mieg J, Thierry-Mieg D, Ashurst J, Jia L, Nakao M, Thomas MA, Mulder N, Karavidopoulou Y, Jin L, Kim S, Yasuda T, Lenhard B, Eveno E, *et al.*: **Integrative annotation of 21,037 human genes validated by full-length cDNA clones.** *PLoS Biol* 2004, 2:e162.
30. Kuhn RM, Karolchik D, Zweig AS, Trumbower H, Thomas DJ, Thakkapallayil A, Sugnet CW, Stanke M, Smith KE, Siepel A, Rosenbloom KR, Rhead B, Raney BJ, Pohl A, Pedersen JS, Hsu F, Hinrichs AS, Harte RA, Diekhans M, Clawson H, Bejerano G, Barber GP, Baertsch R, Haussler D, Kent WJ: **The UCSC genome browser database: update 2007.** *Nucleic Acids Res* 2007, 35:D668-673.
31. Li WH: **Molecular Evolution.** Sinauer Associates, Sunderland, MA; 1997.
32. Kapranov P, Drenkow J, Cheng J, Long J, Helt G, Dike S, Gingeras TR: **Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays.** *Genome Res* 2005, 15:987-997.
33. Liu WM, Mei R, Di X, Ryder TB, Hubbell E, Dee S, Webster TA, Harrington CA, Ho MH, Baid J, Smeekens SP: **Analysis of high density expression microarrays with signed-rank call algorithms.** *Bioinformatics* 2002, 18:1593-1599.
34. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, Gerstein M, Snyder M: **Global identification of human transcribed sequences with genome tiling arrays.** *Science* 2004, 306:2242-2246.
35. Kampa D, Cheng J, Kapranov P, Yamanaka M, Brubaker S, Cawley S, Drenkow J, Piccolboni A, Bekiranov S, Helt G, Tammana H, Gingeras TR: **Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22.** *Genome Res* 2004, 14:331-342.
36. Khaitovich P, Kelso J, Franz H, Visagie J, Giger T, Joerchel S, Petzold E, Green RE, Lachmann M, Paabo S: **Functionality of intergenic transcription: an evolutionary comparison.** *PLoS Genet* 2006, 2:e171.
37. Kondrashov FA, Koonin EV: **Evolution of alternative splicing: deletions, insertions and origin of functional parts of proteins from intron sequences.** *Trends Genet* 2003, 19:115-119.
38. Kimura K, Wakamatsu A, Suzuki Y, Ota T, Nishikawa T, Yamashita R, Yamamoto J, Sekine M, Tsuritani K, Wakaguri H, Ishii S, Sugiyama T, Saito K, Isono Y, Irie R, Kushida N, Yoneyama T, Otsuka R, Kanda K, Yokoi T, Kondo H, Wagatsuma M, Murakawa K, Ishida S, Ishibashi T, Takahashi-Fujii A, Tanase T, Nagai K, Kikuchi H, Nakai K, *et al.*: **Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes.** *Genome Res* 2006, 16:55-65.
39. Xing Y, Lee C: **Alternative splicing and RNA selection pressure – evolutionary consequences for eukaryotic genomes.** *Nat Rev Genet* 2006, 7:499-509.
40. Zhang J, Nielsen R, Yang Z: **Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level.** *Mol Biol Evol* 2005, 22:2472-2479.
41. Bakewell MA, Shi P, Zhang J: **More genes underwent positive selection in chimpanzee evolution than in human evolution.** *Proc Natl Acad Sci USA* 2007, 104:7489-7494.
42. Stewart CB, Disotell TR: **Primate evolution – in and out of Africa.** *Curr Biol* 1998, 8:R582-588.
43. Kaessmann H, Wiebe V, Paabo S: **Extensive nuclear DNA sequence diversity among chimpanzees.** *Science* 1999, 286:1159-1162.
44. Yi S, Ellsworth DL, Li WH: **Slow molecular clocks in Old World monkeys, apes, and humans.** *Mol Biol Evol* 2002, 19:2191-2198.
45. Kondrashov AS: **Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases.** *Hum Mutat* 2003, 21:12-27.
46. Chen WH, Wang XX, Lin W, He XW, Wu ZQ, Lin Y, Hu SN, Wang XN: **Analysis of 10,000 ESTs from lymphocytes of the cynomolgus monkey to improve our understanding of its immune system.** *BMC Genomics* 2006, 7:82.
47. Wallace JC, Korth MJ, Paeper B, Proll SC, Thomas MJ, Magness CL, Iadonato SP, Nelson C, Katze MG: **High-density rhesus macaque oligonucleotide microarray design using early-stage rhesus genome sequence information and human genome annotations.** *BMC Genomics* 2007, 8:28.
48. Suzuki Y, Sugano S: **Construction of a full-length enriched and a 5'-end enriched cDNA library using the oligo-capping method.** *Methods Mol Biol* 2003, 221:73-91.
49. Jurka J: **Repbase update: a database and an electronic journal of repetitive elements.** *Trends Genet* 2000, 16:418-420.
50. Pruitt KD, Tatusova T, Maglott DR: **NCBI Reference Sequence project: update and current status.** *Nucleic Acids Res* 2003, 31:34-37.
51. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, 22:4673-4680.
52. Li WH: **Unbiased estimation of the rates of synonymous and nonsynonymous substitution.** *J Mol Evol* 1993, 36:96-99.
53. Pamilo P, Bianchi NO: **Evolution of the Zfx and Zfy genes: rates and interdependence between the genes.** *Mol Biol Evol* 1993, 10:271-281.
54. Yang Z: **PAML: a program package for phylogenetic analysis by maximum likelihood.** *Comput Appl Biosci* 1997, 13:555-556.
55. Kimura M: **A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences.** *J Mol Evol* 1980, 16:111-120.

ELSEVIER

Journal of Hepatology 48 (2008) 43–50

Journal of Hepatology

www.elsevier.com/locate/jhep

Impact of hepatitis B virus (HBV) X gene integration in liver tissue on hepatocellular carcinoma development in serologically HBV-negative chronic hepatitis C patients☆

Hidenori Toyoda[1], Takashi Kumada[1], Yuji Kaneoka[2], Yoshiki Murakami[3,*,†]

[1]*Department of Gastroenterology, Ogaki Municipal Hospital, 4-86 Minaminokawa, Ogaki, Gifu 503-8502, Japan*
[2]*Department of Surgery, Ogaki Municipal Hospital, 4-86 Minaminokawa, Ogaki, Gifu 503-8502, Japan*
[3]*Laboratory of Human Tumor Virus, Institute for Viral Research, Kyoto University, Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan*

***Background/Aims*: We analyzed hepatitis B virus (HBV) X gene integration in hepatocytes of HBV-negative, chronic hepatitis C (CH-C) patients with mild fibrosis, and prospectively followed these patients for the development of hepatocellular carcinoma (HCC).**

***Methods*: The study included 39 HBV-negative CH-C patients with mild fibrosis. HBV-X integration was determined by Alu-PCR analysis of liver specimens obtained by fine-needle biopsy.**

***Results*: Integration of HBV-X gene sequence into liver genome occurred in 9 of the 39 patients. Six of the 39 patients developed HCC during the 12-year follow-up period. No significant difference was found in the incidence of HCC between patients with and without HBV-X integration. However, the two patients with HBV-X integration who developed HCC did not have cirrhosis at the time when HCC was diagnosed, whereas the four patients without HBV-X integration who developed HCC did have cirrhosis.**

***Conclusions*: Our findings suggest that HBV-X integration detected at the mild fibrosis stage might not indicate a high risk for HCC. HBV-X integration may be associated with HCC development in the absence of cirrhosis. However, we did not find evidence that HBV-X integration directly plays a role in hepatocarcinogenesis in CH-C patients. Further studies will be needed to clarify this point.**

Keywords: **HBV-X integration; Chronic hepatitis C; Hepatocellular carcinoma**

1. Introduction

Chronic viral hepatitis is a leading cause of hepatocellular carcinoma (HCC) worldwide [1–4]. Occult hepatitis B virus (HBV) infection, characterized by the absence of circulating HBV surface antigen [HBsAg] but presence of the HBV genome in serum or liver tissue, has been identified in hepatitis C virus (HCV)-infected patients. HBV may affect the clinical course of chronic hepatitis C (CH-C) [5] and increase the risk of hepatocarcinogenesis [6]. Pollicino reported that both integrated and free HBV–DNA sequences were highly prevalent in the liver tissue of CH-C patients with HCC compared to CH-C patients without HCC [7].

Received 28 March 2007; received in revised form 5 August 2007; accepted 8 August 2007; available online 24 October 2007
Associate Editor: K. Koike

☆ The authors declare that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

* Corresponding author. Tel.: +81 75 751 4034; fax: +81 75 751 3998.
E-mail addresses: ymurakam@virus.kyoto-u.ac.jp, ymurakami@genome.med.kyoto-u.ac.jp (Y. Murakami).
† Present address: Unit of Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Yoshida-Konoe cho, Sakyo-ku, Kyoto 606-8501, Japan. Tel.: +81 75 753 9313; fax: +81 75 753 9314.

0168-8278/$32.00
doi:10.1016/j.jhep.2007.08.016

These observations support the clinical importance of occult HBV as a carcinogenic factor in HBsAg-negative patients with CH-C. However, it remains controversial whether occult HBV increases the risk of HCC in this population [8].

Several studies have investigated the association between HBV integration and HCC in patients with both chronic HCV infection and HCC [8–10]. However, no study has prospectively evaluated whether HBV integration in liver tissue correlates with HCC development in CH-C patients. In a prospective 12-year study, we attempted to clarify whether HBV integration promotes hepatocarcinogenesis in CH-C patients.

2. Materials and methods

2.1. Patients

A total of 67 HBsAg-negative, CH-C patients underwent ultrasonography (US)-guided fine-needle liver biopsy for histological evaluation between January and December 1994. Of these patients, 39 had chronic hepatitis with mild fibrosis (METAVIR score of F0 or F1) [11] and were included in the study. Clinical characteristics of these patients are summarized in Table 1. The patient group contained 30 men and 9 women with a mean age of 49.0 ± 7.6 years. All patients were negative for both serum HBsAg and HBV–DNA and were shown to have persistent HCV infection by nested reverse transcription-polymerase chain reaction (PCR). Sixteen of thirty-nine patients had a history of blood transfusion. No patient had a history of intravenous drug use, tattooing, or acupuncture. No patient had a history of acute hepatitis B. All patients were followed from the time of liver biopsy until October 2006. They underwent periodic US examination and analysis for HCC tumor markers, including α-fetoprotein and des-γ-carboxy prothrombin every 6 months. When a suspicious liver lesion was detected by US or a tumor marker was elevated, the patient underwent further examination by imaging such as computed tomography (CT), magnetic resonance imaging, or angiography. HCC was diagnosed on the basis of typical imaging findings, which include a mosaic pattern with a halo on B-mode US images, hypervascularity on angiographic images, or a high-density mass on arterial-phase dynamic CT images with a low-density mass on portal-phase dynamic CT images obtained with a helical or multidetector raw CT scanner. All patients who developed HCC underwent a hepatectomy; all tumors were less than 3 cm in diameter when detected under this surveillance. The final diagnosis of HCC was based on histologic examination of the tumor tissue taken from resected specimens.

The study protocol conformed to the ethics guidelines of the Declaration of Helsinki (1975). All patients provided written informed consent for analysis of the biopsy specimens, and the Hospital Ethics Committee approved the study.

Table 1
Clinical characteristics of the study patients ($n = 39$)

Age (years)	49.0 ± 7.6
Sex (female/male)	9(23.1)/30(76.9)#
History of blood transfusion	15 (38.5)
Presumed duration of HCV carriage*	19.0 (5–33)##
Alanine aminotransferase (IU/L)	60.1 ± 31.4
Aspartate aminotransferase (IU/L)	45.0 ± 23.8
Gamma glutamyl transpeptidase (mg/L)	51.2 ± 55.3
Albumin (g/dL)	4.11 ± 0.33
Total-bilirubin (mg/dL)	0.74 ± 0.33
Platelet count ($\times 10^4$/ml)	17.9 ± 6.5
HCV RNA concentration ($\times 10^3$IU/mL)	570 (3–4900)##
HCV genotype	
1b	25(64.1)#
2a	11 (28.2)#
2b	3 (7.7)#
HBV surface antigen	0
HBV surface antibody	6(15.4)#
HBV core antibody	25(64.1)#
Fibrosis stage**	
F0	14 (35.9)#
F1	25(64.1)#

HBV, hepatitis B virus; HCV, hepatitis C virus.
* In patients with a history of blood transfusion.
** According to METAVIR score.
\# Percentages are shown in parentheses.
\#\# Median; ranges are shown in parentheses.

2.2. Sample preparation

DNA was extracted from liver tissues obtained at liver biopsy on 1994 with a DNeasy Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. All samples were stored at −80 °C and carefully handled to avoid contamination with nucleic acids.

2.3. Detection of viral–host junctions

A PCR-based technique, Alu-PCR, one of the most effective procedures to detect junctions between integrated HBV–DNA and human DNA, was used to amplify viral–host junctions using 100 ng of genomic DNA [12–14] (Table 2). The sensitivity study for this PCR was performed using human hepatoma cell line Huh-2 cells that contain 1 copy per cell of integrated HBV (kindly provided by Dr. K Koike from Department of Gene Research, Cancer Institute, Tokyo) [15]. Amplified PCR products were analyzed by electrophoresis on 1.0% agarose gel and transferred to a Hybond-N^+ nylon membrane (Amersham Pharmacia, Buckinghamshire, UK). About 3.2 kb of the HBV X genome (HBV-X) was amplified according to the method of Günther et al. [16]. HBV-specific bands were then detected by hybridization with a DIG labeled HBV probe (Roche, Mannheim, Germany).

2.4. Direct sequencing

The amplified viral/host junctions were purified with an Easy Trap Kit (Takara, Otsu, Japan) and sequenced using a Prism Taq DyeDeoxy Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA), according to the manufacturer's instructions. Products were precipitated with ethanol and analyzed with a 377 Prism DNA Sequencer (Applied Biosystems Inc.). To identify the HBV-X integration site, we used BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) to compare sequences adjacent to the integrated HBV–DNA with the human genome.

2.5. Other serological and virological tests

HBV surface antigen, surface antibody, and core antibody were measured with ARCHITECT HBsAg QT, ARCHITECT anti-HBs, and ARCHITECT anti-HBc, respectively (Abbott Japan, Tokyo, Japan). Serum HBV–DNA was measured by the Amplicor HBV test (detection limit, 400 copies/mL; Roche Diagnostics, Branchburg, NJ). HCV genotype was determined by PCR with genotype-specific primers [17,18]. HCV RNA concentration was measured by a quantitative PCR assay (detection limit, 5000 copies/mL; Amplicor GT-HCV Monitor, Version 2.0; Roche Molecular Systems, Pleasanton, CA).

2.6. Statistical analyses

Data are expressed as means ± SD or the median and range. Differences in the proportion of patients with and without HBV-X integration were analyzed by χ^2 test. Differences in quantitative values were analyzed by Mann–Whitney U test. For the incidence of HCC

Table 2
Sequences of primers for detection of viral–host junctions

Primer name	Primer sequence	HBV portion and note
UP5	5′-CAGUGCCAAGUGUUUGCUGACGCCAAAGUGCUGGGAUUA-3′	Alu-sense
T3-515	5′-AUUAACCCUCACUAAAGCCUCGAUAGAUYRYRCCAYUGCAC-3′	Alu-antisense
UP6	5′-CAAGTGTTTGCTGACGCCAAAG-3′	Alu-sense (tag)
midT3	5′-ATTAACCCTCACTAAAGCCTCG-3′	Alu-antisense (tag)
pUTP	5′-ACAUGAACCUUUACCCCGUUGC-3′	1131–1152 HB1 (HBV-X)
MD37	5′-TGCCAAGTGTTTGCTGACGC-3′	1174–1193 HB2 (HBV-X)
MD60	5′-CTGCCGATCCATACTGCGGAAC-3′	1258–1279 HB3 (HBV-X)

Numbering of nucleotides is according to Ono et al. [31]. U = dUTP; Y = (C,T); R = (A,G).

development, the date of the initial liver biopsy was defined as time zero. Data pertaining to patients who did not develop HCC were censored. The Kaplan–Meier method was used to calculate the incidence of HCC, and the log-rank test was used to analyze differences. The JMP statistical software package, version 4.0, (SAS Institute, Cary, NC) was used for all statistical analyses. All p values were derived from two-tailed tests, and $p < 0.05$ was considered statistically significant.

3. Results

3.1. Integration of hepatitis B viral genome and patient characteristics

The sensitivity of the PCR amplification was first determined with hepatoma cell line Huh-2 cells. When we made a tenfold serial dilution of Huh-2 cells with normal human PBMC without a history of liver disease, we could detect viral–host junctions at about 100 copies per reaction by the PCR (Fig. 1a).

We amplified virus–host DNA junctions from the liver of CH-C patients and detected several bands on 1.0%-agarose gels (Fig. 1b). Sequencing these PCR products revealed HBV-X integration in 9 of the 39 (23.1%) patients. Nineteen viral–host junctions were detected in these 9 patients. In 4 of these 9 patients, multiple integration sites (range, 2–6) were present. For example, 6 viral–host junctions were detected in patient 15, and the adjacent host sequences were from 6 different chromosomes (red circle, Fig. 2). In the other 5 cases, a single integration site was detected. The sites of HBV-X integration are shown in Fig. 2.

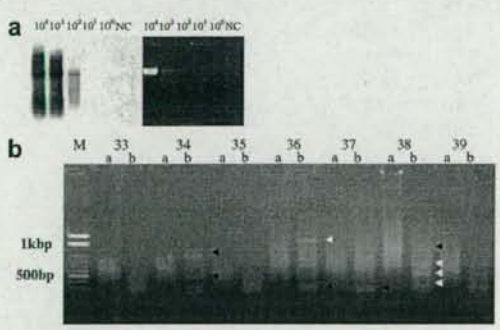


Fig. 1. The detection of HBV-X–host junction by Alu-PCR analysis. (a) The sensitivity study of Alu-PCR method. Serially diluted genomic DNA contained with HBV integrant was amplified by using HBV-X and Alu antisense primer pair. Left is Southern blot analysis from the gel electrophoresis (right). (b) The numbers indicate the individual patients, and a and b indicate the primer pair used for amplification (a, HBV-X primer and Alu sense; b, HBV-X primer and Alu antisense). The PCR strategy and the primer sequences used in this study were previously described [12–14]. Arrowheads indicate PCR products with HBV-X–host junctional sequences (white) and without HBV-X–host junctions (black).

Clinical characteristics of patients with and without HBV-X integration are summarized in Table 3. There were no differences in the clinical characteristics. During the observation period, 4 of 9 (44.4%) patients with HBV-X integration and 13 of 30 (43.3%) patients without HBV-X integration received interferon monotherapy. These percentages did not differ significantly.

3.2. Host genome sequences at sites of HBV-X integration

The sites of host integration were divided into two groups: (1) genes already known and/or fully characterized but not previously shown to be involved in carcinogenesis (1 integration site; T cell lymphoma invasion and metastasis 1 [TIAM1] in Patient 8), and (2) unknown open reading frames (ORFs) or genes belonging to a known gene family but not functionally characterized (18 integration sites). The HBV genome ORF was integrated in both the same and opposite orientations of the host gene and both proximal to and into host genes (Table 4).

3.3. Development of HCC

Over the 12-year follow-up period, HCC developed in 6 of the 39 (15.4%) patients. HCC developed in 4 of the 30 (13.3%) patients without HBV-X integration and in 2 of the 9 (22.2%) patients with HBV-X integration (Fig. 3). The difference in the incidence of HCC between patients with and without HBV-X integration was not significant ($p = 0.8041$). Patient age, sex, and histologic data at the time of HBV-X integration analysis and at the time of HCC diagnosis are shown in Table 5. All patients who developed HCC were males. Age at the time HCC developed did not differ between patients

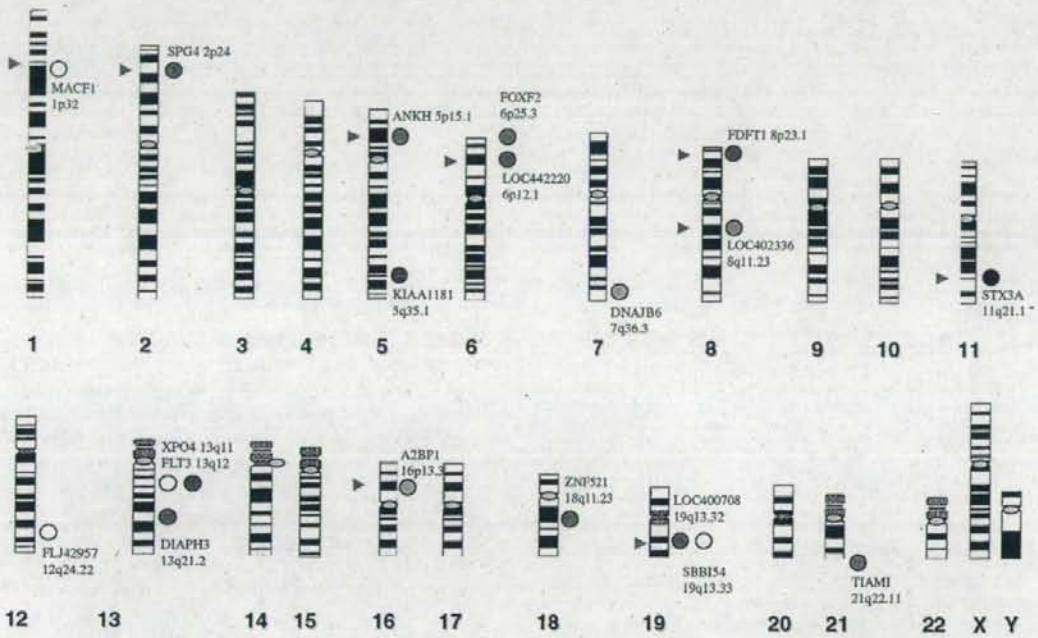


Fig. 2. Chromosomal distribution of HBV-X integration sites. Circles indicate viral integration sites, and the circle color denotes the sample. For example, the three white spots indicate three viral integration sites detected in the same specimen. Gene names and chromosomal localizations are also noted. Red arrowheads indicate DNA fragile sites [32].

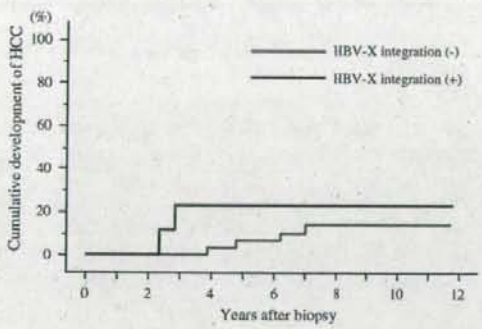


Fig. 3. Kaplan–Meier curves for the incidence of hepatocellular carcinoma (HCC). The blue and red lines represent the incidence of HCC in patients with and without HBV-X integration, respectively. No significant difference was observed between the two groups ($p = 0.8041$).

with and without HBV-X integration. Five of 6 patients who developed HCC (except for patient No. 34) had received interferon therapy, but all of them remained HCV positive. All 4 patients without HBV-X integration who developed HCC had cirrhosis at the time HCC was diagnosed. In contrast, the fibrosis stage was moderate or mild (F1 or F2) in the 2 patients with HBV-X integration who developed HCC. No patient was positive for the circulating low-level HBV–DNA analyzed with a highly sensitive HBV–DNA detection method (detection limit, 35 copies/mL; COBAS Taq-Man HBV test, Roche Diagnostics) at the time of HCC diagnosis [19].

We attempted to detect HBV–host junction by the same Alu-PCR method in resected HCC materials that developed in 4 patients (patients #9, 21, 34, and 38) using paraffin-embedded samples. HBV-X integration was detected in HCC materials of none of 4 patients (data not shown).

4. Discussion

This is the first prospective study to analyze HBV integration into the host hepatocyte genome of CH-C patients with mild fibrosis and then to follow these patients over a long period for the development of HCC. Previous studies investigated HBV integration in HCC tissue of patients chronically infected with HCV [8–10] or in HCC tissue of patients without hepatitis virus infection [20]. However, in these studies, HBV integration was analyzed in cancerous and non-cancerous tissue after the development of HCC, and thus the effect of HBV integration on the development of HCC in CH-C patients was not investigated.

Table 3
Characteristics of patients with and without HBV-X–DNA integration

	HBV-X–DNA integration (−) ($n = 30$)	HBV-X–DNA integration (+) ($n = 9$)
Age (years)	48.9 ± 7.6	49.6 ± 7.7
Sex (female/male)	6 (20.0)/24 (80.0)[#]	3 (33.3)/6 (66.7)[#]
History of blood transfusion	11(36.7)[#]	4 (44.4)[#]
Presumed duration of HCV carriage[*]	19.0 (5–33)[##]	22.5 (12–33)[##]
Alanine aminotransferase (IU/L)	60.0 ± 31.8	60.4 ± 31.8
Aspartate aminotransferase (IU/L)	46.2 ± 25.9	41.0 ± 15.8
Gamma glutamyl transpeptidase (mg/L)	49.5 ± 39.0	34.6 ± 29.2
Albumin (g/dL)	4.08 ± 0.37	4.22 ± 0.14
Total-bilirubin (mg/dL)	0.70 ± 0.33	0.84 ± 0.33
Platelet count ($\times 10^4$/ml)	18.0 ± 5.3	17.6 ± 9.9
HCV RNA concentration ($\times 10^3$IU/mL)	790 (3–4900)[##]	320 (3–2100)[##]
HCV genotype		
1b	19(63.3)[#]	6 (66.7)[#]
2a	8 (26.7)[#]	3 (33.3)[#]
2b	3 (10.0)[#]	0
HBs antibody (+)	4(13.3)[#]	2 (22.2)[#]
HBc antibody (+)	20 (66.7)[#]	5 (55.6)[#]
Fibrosis stage[**]		
F0	10 (33.3)[#]	4 (44.4)[#]
F1	20 (66.7)[#]	5 (55.6)[#]

HBV, hepatitis B virus; HCV, hepatitis C virus.
[*] In patients with a history of blood transfusion.
[**] According to METAVIR score.
[#] Percentages are shown in parentheses.
[##] Median; ranges are shown in parentheses.

Table 4
Genes and sequences of HBV-X–DNA integration sites

No.	Supercontig	Position	Orientation	Chromosomal localization	Name	Location	Name/function
8.	NT034880	1375087	Same	6p25.3	FOXF2	39 kb upstream	Forkhead box F2
8.	NT086666	14245273	Opposite	5p15.1	ANKH	177 kb upstream	Ankylosis, progressive homolog
8.	NT011512	18351760	Same	21q22.11	TIAM1	21.5 kb upstream	T-cell lymphoma invasion and metastasis 1
15.	NT_022184	11183657	Same	2p24	SPG4	Intronic seq	Spastic paraplegia 4 (autosomal dominant; spastin)
15.	NT_024524	41428064	Same	13q21.2	DIAPH3	Intronic seq	Diaphanous homolog 3 (*Drosophila*)
15.	NT011109	19337933	Same	19q13.32	LOC400708	3.1 kb upstream	Similar to Serine/threonine protein phosphatase 5 (PP5)
15.	NT_077531	4155242	Opposite	8p23.1	FDFT1	Intronic seq	Farnesyl-diphosphate farnesyltransferase 1
15.	NT_010966	4345775	Opposite	18q11.2	ZNF521	23 kb upstream	Zinc finger protein 521
15.	NT_007592	46424722	Same	6p12.1	LOC442220	5.3 kb upstream	Similar to nitrogen fixation cluster-like
21.	NT_023133	17103986	Opposite	5q35.1	KIAA1181	38 kb downstream	Endoplasmic reticulum-golgi intermediate compartment 32 kDa protein
22.	NT011109	23275592	Same	19q13.33	SBBI54	Intronic seq	Hypothetical transmembrane protein SBBI54
23.	NT008183	6327145	Opposite	8q11.23	LOC402336	16.9 kb upstream	Similar to L21 ribosomal protein
24.	NT_024524	2436145	Opposite	13q11	XPO4	12.6 kb upstream	Exportin 4
27.	NT_007741	2000247	Opposite	7q36.3	DNAJB6	4 kb downstream	DnaJ (Hsp40) homolog, subfamily B, member 6 Homo sapiens
27.	NT086834	6475804	Opposite	16p13.3	A2BP1	31.9 kb upstream	Ataxin 2-binding protein 1
36.	NT_033903	4799121	Opposite	11q21.1	STX3A	29 kb downstream	Syntaxin3A
38.	NT_009775	7468765	Opposite	12q24.22	FLJ42957	71 kb downstream	FLJ42957 protein
38.	NT_024524	9545675	Opposite	13q12	FLT3	20 kb downstream	Fms-related tyrosine kinase 3
38.	NT004511	9911738	Opposite	1p32	MACF1	Intronic seq	Microtubule-actin crosslinking factor 1

In three studies of HCV-related HCC, the rates of HBV integration in tumor tissue are discrepant: 55.6% (10 out of 18 cases) [8], 29.4% (10 out of 34 cases) [10], and 0% (0 out of 21 cases) [9]. Clonal expansion of hepatocytes containing integrated HBV in association with cancer progression may increase the detection rate of HBV integration. Conversely, clonal expansion of cancerous hepatocytes without HBV integration may decrease the

Table 5
Cases of HCC development

No.	Sex	Age at biopsy	Fibrosis at biopsy	Interval between biopsy and HCC development	Age at HCC development	Fibrosis at HCC development*	HBV-X–DNA integration
7.	M	61	F1	4y.	65	F4	(−)
9.	M	57	F1	5y.	62	F4	(−)
21.	M	56	F1	3y.	64	F2	(+)
28.	M	56	F1	5y.	61	F4	(−)
34.	M	47	F1	7y.	54	F4	(−)
38.	M	55	F0	2y.	57	F1	(+)

* Non-cancerous tissue.

detection of HBV integration. Therefore, hepatocyte clonal expansion may account for discrepancies in the rates of HBV integration between studies. In contrast, clonal expansion of hepatocytes is unlikely in cases of CH-C with mild fibrosis but without HCC. The prevalence of HBV-X integration in our patient population (23.1%), therefore, represents the actual rate of HBV-X integration in CH-C patients. The number of HBV-X -host integration sites in these patients was smaller than patients with chronic hepatitis B and similar to patients with acute hepatitis B in our previous study with the same detection method for HBV integration [13].

HBV integration is detected in approximately 90% of liver tumor samples from patients with HBsAg [21]. HBV insertional mutagenesis is an important step in many cases of hepatocarcinogenesis in patients with chronic HBV infection. Chromosomal inversions, translocations, or micro deletions can occur at the integration sites, causing tumors to develop in some patients [22,23]. Several tumor-associated genes have been identified adjacent to HBV integration sites [24,25]. However, HBV does not integrate in or near a tumor-associated gene in most HBV-infected individuals. Rather, HBV–DNA integrates randomly into host DNA in HBV-related HCC [21,26,27]. This random integration also appears in patients with HCV-related HCC, although one study suggested that HBV–DNA integrates into tumor-associated genes of some HCC patients without HBsAg [8].

In the present study of CH-C patients without HCC, the HBV-X integration sites were distributed across the genome with little similarity and the host sequences adjacent to the viral genome were divergent. These data are consistent with our previous results on HBV-infected patients with the same detection method for HBV-X integration [14]. In the present study, we did not detect HBV-X integration into genes associated with carcinogenesis. Because HBV–DNA integrates randomly into host DNA and the number of HBV-integration sites was smaller in CH-C patients compared to chronic hepatitis B patients [13], the likelihood of HBV integrating into genes associated with carcinogenesis would be considerably low.

We analyzed HBV-X integration in CH-C patients with mild fibrosis and prospectively observed the patient group for 12 years. There was no statistically significant difference in the incidence of HCC between patients with and without HBV-X integration. Taken together with results from clinical observations and genetic analyses, these data suggest that testing HBV-X integration at a mild fibrosis stage may not predict the likelihood of CH-C patients developing hepatocarcinogenesis. However, the lack of statistical significance in the incidence of HCC could be partly because of the small number of study patients. Future studies with a larger patient population may detect patients with HBV integration in tumor-associated genes and a higher incidence of HCC development in CH-C patients with HBV integration.

In the present study, there was no cirrhosis in non-cancerous liver tissue surrounding the tumor at the time of HCC development, and fibrosis was not severe (stage F1 or F2) in patients with HBV-X integration. In contrast, all 4 HCC patients without HBV-X integration had cirrhosis (stage F4). In addition, the interval between the analysis of HBV-X integration and HCC development was shorter in patients with HBV-X integration than those without HBV-X integration. The stage of fibrosis, especially the presence of cirrhosis, is related closely to the incidence of HCV-related HCC [28], and most patients with HCV-related HCC have cirrhosis [10,29]. Our results showed that HCC develops in the absence of cirrhosis in some CH-C patients with HBV-X integration, and this may suggest the possibility that HBV-X integration may play a role in accelerated hepatocarcinogenesis in some cases. However, we did not detect HBV-X integration in paraffin-embedded resected HCC materials of both 2 patients with HBV-X integration at liver biopsy (patients #21 and #38). Although this can be partly due to the use of paraffin-embedded materials for analyses of integration (unfortunately frozen section was not available), we did not find the evidence that HBV-X integration directly played a role in hepatocarcinogenesis in the present study.

There are several limitations of the study. The detection of HBV integration with PCR using Alu repeats may limit the identification of HBV-X sequence integration sites that are far away from the priming site,

therefore, restricting the sensitivity of the assay as the amplicon size increases. In addition, detection of HBV integration only using the X gene-specific primers makes infeasible identification of integration sites of other virus gene sequences. Further, integrated HBV genome can limit or negate entirely the HBV X primer-binding site, because HBV sequences may be deleted upon integration. The Alu-PCR method used in the present study, therefore, may underestimate the integration of HBV in CH-C patients.

In summary, HBV-X integration was detected in 9 of 39 CH-C patients and the number of HBV-X–host integration sites in these patients was similar to patients with acute hepatitis B. They were distributed across the genome with little similarity. In the prospective observation of CH-C patients over 12 years, HBV-X integration detected at the mild fibrosis stage might not indicate a high risk for HCC during the course of CH-C. Although HBV-X integration may be associated with HCC development in the absence of cirrhosis, we did not find evidence that HBV-X integration directly plays a role for hepatocarcinogenesis in this patient population. Further studies with more sensitive and reliable method than Alu-PCR method for the detection of HBV integration are needed to elucidate the association between HBV integration and HCC development in CH-C patients without cirrhosis. Also, the analyses for HBV integration in frozen sections of resected HCC materials from CH-C patients in whom HBV integration was detected at the mild fibrosis stage may provide the evidence for direct association between HBV integration and accelerated hepatocarcinogenesis in this population. In addition, the association between genotype of integrated HBV and hepatocarcinogenesis in this population should also be investigated in the future, because the potential incidence of HCC reportedly differs according to HBV genotype in case of HBV-infected patients [30].

Acknowledgement

The authors thank Prof. Kunitada Shimotohno, Laboratory of Human Tumor Virus, Institute for Viral Research, Kyoto University, for his advice and comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhep.2007.08.016.

References

[1] Beasley RP. Hepatitis B virus. The major etiology of hepatocellular carcinoma. Cancer 1988;61:1942–1956.

[2] Kiyosawa K, Sodeyama T, Tanaka E, Gibo Y, Yoshizawa K, Nakano Y, et al. Interrelationship of blood transfusion, non-A, non-B hepatitis and hepatocellular carcinoma: analysis by detection of antibody to hepatitis C virus. Hepatology 1990;12:671–675.

[3] Di Bisceglie AM, Goodman ZD, Ishak KG, Hoofnagle JH, Melpolder JJ, Alter HJ. Long-term clinical and histopathological follow-up of chronic posttransfusion hepatitis. Hepatology 1991;14:969–974.

[4] Brechot C, Jaffredo F, Lagorce D, Gerken G, Meyer zum Buschenfelde K, Papakonstontinou A, et al. Impact of HBV, HCV and GBV-C/HGV on hepatocellular carcinomas in Europe: results of a European concerted action. J Hepatol 1998;29:173–183.

[5] Cacciola I, Pollicino T, Squadrito G, Cerenzia G, Orlando ME, Raimondo G. Occult hepatitis B virus infection in patients with chronic hepatitis C liver disease. N Engl J Med 1999;341:22–26.

[6] Tamori A, Nishiguchi S, Kubo S, Koh N, Moriyama Y, Fujimoto S, et al. Possible contribution to hepatocarcinogenesis of X transcript of hepatitis B virus in Japanese patients with hepatitis C virus. Hepatology 1999;29:1429–1434.

[7] Pollicino T, Squadrito G, Cerenzia G, Cacciola I, Raffa G, Craxi A, et al. Hepatitis B virus maintains its pro-oncogenic properties in the case of occult HBV infection. Gastroenterology 2004;126:102–110.

[8] Urashima T, Saigo K, Kobayashi S, Imaseki H, Matsubara H, Koide Y, et al. Identification of hepatitis B virus integration in hepatitis C virus-infected hepatocellular carcinoma tissues. J Hepatol 1997;26:771–778.

[9] Kawai S, Yokosuka O, Imazeki F, Maru Y, Saisho H. State of HBV DNA in HBsAg-negative, anti-HCV-positive hepatocellular carcinoma: existence of HBV DNA possibly as nonintegrated form with analysis by Alu-HBV DNA PCR and conventional HBV PCR. J Med Virol 2001;64:410–418.

[10] Tamori A, Nishiguchi S, Kubo S, Enomoto M, Koh N, Takeda T, et al. Sequencing of human–viral DNA junctions in hepatocellular carcinoma from patients with HCV and occult HBV infection. J Med Virol 2003;69:475–481.

[11] Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology 1994;20:15–20.

[12] Minami M, Poussin K, Brechot C, Paterlini P. A novel PCR technique using Alu specific primers to identify unknown flanking sequences from the human genome. Genomics 1995;29:403–408.

[13] Murakami Y, Minami M, Daimon Y, Okanoue T. Hepatitis B virus DNA in liver, serum, and peripheral blood mononuclear cells after the clearance of serum hepatitis B virus surface antigen. J Med Virol 2004;72:203–214.

[14] Murakami Y, Saigo K, Takashima H, Minami M, Okanoue T, Brechot C, et al. Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. Gut 2005;54:1162–1168.

[15] Koike K, Kobayashi M, Mizusawa H, Yoshida E, Yaginuma K, Taira M. Rearrangement of the surface antigen gene of hepatitis B virus integrated in the human hepatoma cell lines. Nucleic Acids Res 1983;25:5391–5402.

[16] Gunther S, Li BC, Miska S, Kruger DH, Meisel H, Will H. A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients. J Virol 1995;69:5437–5444.

[17] Okamoto H, Kobata S, Tokita H, Inoue T, Woodfield GD, Holland PV, et al. A second-generation method of genotyping hepatitis C virus by the polymerase chain reaction with sense and antisense primers deduced from the core gene. J Virol Methods 1996;57:31–45.

[18] Simmonds P, Alberti A, Alter HJ, Bonino F, Bradley DW, Brechot C, et al. A proposed system for the nomenclature of hepatitis C viral genotypes. Hepatology 1994;19:1321–1324.
[19] Toyoda H, Kumada T, Kiriyama S, Sone Y, Tanikawa M, Hisanaga Y, et al. Prevalence of low-level hepatitis B viremia in patients with HBV surface antigen-negative hepatocellular carcinoma with and without hepatitis C virus infection in Japan: analysis by COBAS TaqMan real-time PCR. Intervirology 2007;50:241–244.
[20] Tamori A, Nishiguchi S, Kubo S, Narimatsu T, Habu D, Takeda T, et al. HBV DNA integration and HBV-transcript expression in non-B, non-C hepatocellular carcinoma in Japan. J Med Virol 2003;71:492–498.
[21] Brechot C, Gozuacik D, Murakami Y, Paterlini-Brechot P. Molecular bases for the development of hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). Semin Cancer Biol 2000;10:211–231.
[22] Hino O, Shows TB, Rogler CE. Hepatitis B virus integration site in hepatocellular carcinoma at chromosome 17;18 translocation. Proc Natl Acad Sci USA 1986;83:8338–8342.
[23] Nakamura T, Tokino T, Nagaya T, Matsubara K. Microdeletion associated with the integration process of hepatitis B virus DNA. Nucleic Acids Res 1988;16:4865–4873.
[24] Dejean A, Bougueleret L, Grzeschik KH, Tiollais P. Hepatitis B virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. Nature 1986;322:70–72.
[25] Wang J, Chenivesse X, Henglein B, Brechot C. Hepatitis B virus integration in a cyclin A gene in a hepatocellular carcinoma. Nature 1990;343:555–557.
[26] Nagaya T, Nakamura T, Tokino T, Tsurimoto T, Imai M, Mayumi T, et al. The mode of hepatitis B virus DNA integration in chromosomes of human hepatocellular carcinoma. Genes Dev 1987;1:773–782.
[27] Takada S, Gotoh Y, Hayashi S, Yoshida M, Koike K. Structural rearrangement of integrated hepatitis B virus DNA as well as cellular flanking DNA is present in chronically infected hepatic tissues. J Virol 1990;64:822–828.
[28] Yoshida H, Shiratori Y, Moriyama M, Arakawa Y, Ide T, Sata M, et al. Interferon therapy reduces the risk for hepatocellular carcinoma: national surveillance program of cirrhotic and noncirrhotic patients with chronic hepatitis C in Japan. Ann Int Med 1999;131:174–181.
[29] Tamori A, Nishiguchi S, Shiomi S, Hayashi T, Kobayashi S, Habu D, et al. Hepatitis B virus DNA integration in hepatocellular carcinoma after interferon-induced disappearance of hepatitis C virus. Am J Gastroenterol 2005;100:1748–1753.
[30] Chan HL, Hui AY, Wong ML, Tse AM, Hung LC, Wong VW, et al. Genotype C hepatitis B virus infection is associated with an increased risk of hepatocellular carcinoma. Gut 2004;53:1494–1498.
[31] Ono Y, Onda H, Sasada H, Igarashi K, Sugino Y, Nishioka K. The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adw. Nucleic Acid Res 1983;11:1747–1757.
[32] Kusano N, Okita K, Shirahashi H, Harada T, Shiraishi K, Oga A, et al. Chromosomal imbalances detected by comparative genomic hybridization are associated with outcome of patients with hepatocellular carcinoma. Cancer 2002;94:746–751.

HUMAN MUTATION 29(5), 703–708, 2008

RESEARCH ARTICLE

Integration of Hepatitis B Virus DNA Into the Myeloid/Lymphoid or Mixed-Lineage Leukemia (*MLL4*) Gene and Rearrangements of *MLL4* in Human Hepatocellular Carcinoma

Kenichi Saigo,[1,2] Kenichi Yoshida,[3]* Ryuji Ikeda,[3] Yoshiko Sakamoto,[3] Yoshiki Murakami,[4] Tetsuro Urashima,[1] Takehide Asano,[5] Takashi Kenmochi,[2] and Ituro Inoue[3,6]

[1]*Second Department of Surgery, School of Medicine, Chiba University, Chiba, Japan;* [2]*Clinical Research Center, Chiba-East National Hospital, Chiba, Japan;* [3]*Division of Genetic Diagnosis, Institute of Medical Science, University of Tokyo, Tokyo, Japan;* [4]*Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan;* [5]*Hepato-Pancreato-Biliary Surgery, School of Medicine, Teikyo University, Tokyo, Japan;* [6]*Division of Molecular Life Science, School of Medicine, Tokai University, Kanagawa, Japan*

Communicated by Dr. Haig H. Kazazian, Jr.

Integration of hepatitis B virus (HBV) DNA into host DNA is detected in about 90% of HBV-related hepatocellular carcinoma (HCC), but the preferential sites of the viral integration etiologically relevant to oncogenesis have been controversial. By using an adaptor-ligation/suppression-PCR, we identified four integrations into the myeloid/lymphoid or mixed-lineage leukemia 4 (*MLL4*) gene from 10 HCC patients with positive HBV surface antigen (HBsAg). Determination of the cellular-virus DNA junction demonstrated that various lengths of the virus were integrated within 300 bp of intron 3 flanked by the Alu element of *MLL4*. Chimeric hepatitis B virus X gene (HBx)/*MLL4* transcripts and the HBx fusion proteins were detected. DNA microarray revealed that HBx/MLL4 fusion proteins suppressed unique genes in HepG2 cells. Finally, chromosomal translocations of intron 3 of *MLL4* to the specific region of chromosome 17p11.2 in 22 out of 32 HCC patients were observed, showing that the intron 3 region of *MLL4* gene would be a target of translocation breakpoint. In conclusion, the present data suggest that the translocation breakpoint of *MLL4* gene is one of the preferential targets for HBV DNA integration into the *MLL4* gene and the HBV DNA integration may be involved in liver oncogenesis. Hum Mutat 29(5), 703–708, 2008. © 2008 Wiley-Liss, Inc.

KEY WORDS: hepatocellular carcinoma; DNA integration; hepatitis B virus; HBx; *MLL4*

INTRODUCTION

Chronic human hepatitis B virus (HBV) infection causes mild to severe liver diseases, such as chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma (HCC) [Block et al., 2003]. Nearly 25% of patients with chronic HBV infections terminate in untreatable liver cancer. HBV DNA frequently integrates into the human host genome, whereby insertional mutagenesis plays a crucial role in oncogenesis [Brechot et al., 2000; Gozuacik et al., 2001]. Although integration of HBV DNA is thought to be involved in oncogenesis of human hepatocytes, preferential HBV DNA integration sites targeting cellular genes were not identified until recently. Two groups have reported that HBV DNA is preferentially integrated into the human telomerase reverse transcriptase (*TERT*) gene (MIM# 187270) in HCC [Ferber et al., 2003; Paterlini-Brechot et al., 2003].

In this study, we investigated the integrated HBV DNA and flanking cellular DNA sequences. In four cases, integrations of HBV DNA into intron 3 of the myeloid/lymphoid or mixed-lineage leukemia 4 (*MLL4*) gene (MIM# 606834) were demonstrated, indicating that *MLL4* serves as a cellular target for HBV in liver oncogenesis. The *MLL4* gene is a human member of the *MLL* gene family locating on chromosome 19q13.1 [FitzGerald and Diaz, 1999], where a frequent rearrangement or amplification has been reported in solid tumors [Mitelman et al., 1997; Curtis et al., 1998]. Subsequently, we detected chromosomal translocation between intron 3 of *MLL4* and a specific region of chromosome17p11.2 in 22 HCC samples. These results indicate that intron 3 of the *MLL4* gene is one of the sites of translocation break point, which serves as a preferential target for HBV DNA integration, and may be implicated in the etiology of liver oncogenesis.

The Supplementary Material referred to in this article can be accessed at http://www.interscience.wiley.com/jpages/1059-7794/suppmat.

Received 18 September 2007; accepted revised manuscript 16 November 2007.

*Correspondence to: Kenichi Yoshida, Department of Life Sciences, Meiji University School of Agriculture, 1-1-1 Higashimita, Tama-ku, Kawasaki, Kanagawa 214-8571, Japan (present address). E-mail: yoshida@isc.meiji.ac.jp

Grant sponsor: CREST of Japan Science and Technology (II).
DOI 10.1002/humu.20701
Published online 4 March 2008 in Wiley InterScience (www.interscience.wiley.com).

PATIENTS AND METHODS

Patients

We studied 32 Japanese patients with HCC who had undergone hepatic resection without preoperative therapies at The Second Department of Surgery, Chiba University Hospital between 1987 and 2003. Serological tests for HBV were done by EIA kit (Dainabot, Tokyo, Japan) for HBV surface antigen (HBsAg), and RIA kits (Dainabot) for anti-HBs and anti-HBc antibodies. Anti-HCV antibody was tested by a recombinant immunoblot assay (Ortho Diagnostic, Westwood, MA). The study protocol conformed to the ethical guidelines of the Declaration of Helsinki (1975) and was approved by the Institutional Review Board (IRB) of Chiba University, School of Medicine. All patients gave written informed consent.

PCR and Southern Blot

HBV/cellular DNA junctions in the tumor tissues were analyzed by an adaptor-ligation/suppression-PCR [Siebert et al., 1995], according to Genomewalker Kits (Clonetech, Mountain View, CA) (Supplementary Fig. S1, available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat). Primer sequences used for PCR detection of HBV/*MLL4* and *MLL4*/HBV junctions and chromosome 19/chromosome 17 boundaries are listed in Supplementary Table S1.

Hind III-digested DNA (10 μg) were electrophoresed on 1.0% agarose gel and blotted onto nylon membrane (Hybond N+; GE Healthcare, Buckinghamshire, UK). The membrane was first hybridized with ^{32}P-labeled hepatitis B virus X gene (HBx) probe and the blot was autoradiographed. After dehybridization of the same membrane, a rehybridization was carried out with ^{32}P-labeled *MLL4* probe (the PCR products spanning exon 4 and exon 5), and autoradiographed.

RT-PCR

Total cellular RNA was extracted using Trizol (Invitrogen, Carlsbad, CA). An RT-PCR was performed with SuperScript One-Step RT-PCR system (Invitrogen) with gene-specific primers on exon 5 and exon 6 of *MLL4*. MD26c primer was used as the common sense primer. The PCR products were subjected to sequencing analyses.

Immunoprecipitation and Western Blot

Tumor tissues were lysed in a buffer containing 0.1% SDS, 0.5% deoxycholate, 1% NP-40, 150 mM NaCl, 50 mM Tris-HCl (pH 8.0), protease inhibitors (complete protease inhibitor tablets; Roche, Basel, Switzerland), and centrifuged. The supernatant was incubated with anti-HBx monoclonal antibody, generously provided by Dr. Yosef Shaul (Weizman Institute of Science), and immunoprecipitation/Western blot was performed with a standard protocol. Anti-Flag antibody was purchased from Sigma-Aldrich (St. Louis, MO).

HBx/*MLL4* Expression Plasmid, Transfection, and DNA Microarray

The HBx expression vector, pECE-X, was a gift of Dr. Jinghsiung James Ou (University of Southern California). Human *MLL4* partial cDNA clone KIAA0304 (accession number AB002302.2) was obtained from Kazusa DNA Research Institute (Chiba, Japan). We deleted intron 7 sequence from KIAA0304 and constructed N-terminally Flag-tagged HBx/*MLL4* chimeric sequence in pcDNA3 (Invitrogen). Human hepatoma cell line HepG2 (RCB1648; RIKEN Cell Bank, Tsukuba, Japan) were transfected using Lipofectamine (Invitrogen). After 48 hr of transfection, total RNA was recovered and the microarray analysis including 12,814 unique clones from Incyte UniGene 1 was performed according to the manufacturer's instructions (Agilent Technologies, Santa Clara, CA).

RESULTS

Detection and Sequence Analysis of HBV/Cellular DNA Junctions

A total of 10 tumor specimens from HCC patients with positive HBsAg were examined for HBV DNA (accession number AB033550.1) integrations into cellular genome. The clinical backgrounds of the patients are summarized (Supplementary Table S2). We could detect four integrations into the *MLL4* gene on chromosome 19q13.1 and one into the *TERT* gene (Table 1). Integration sites of *MLL4* (accession number AD000671.1) from the four patients were all in intron 3 of the *MLL4* gene (Fig. 1A; Table 1) within or flanked with the Alu repeat sequence (Fig. 1B). As shown in Fig. 1C, full-length viral integration could be expected in HCC131 (g.17752_17753insAB033550.1:g.1827_1826), while truncated virus integrations were detected in the other three tissues, HCC143 (g.17817_17818insAB033550.1:g.2974_1794), HCC146 (g.17514_17515insAB033550.1:g.?_1807), and HCC002 (g.17542_17543insAB033550.1:g.1051_1762). In all four patients, the viral junctions described above were located in the vicinity of DR1, suggesting that the DR1 region is the preferred viral junction for HBV DNA integration.

On Southern blot analysis, clonally integrated HBV DNA sequences were detected in the tumor tissue of HCC131 and a positive control. We encountered the limitations with the heterogeneity of other samples. Using Southern blot hybridization

TABLE 1. Detection of HBV Integration and the Translocation of *MLL4* in HCC

Case no.	Chromosome	Accession no.	Gene	t(17;19)(p11;q13.1)
HCC131	19q13.1	AD000671.1	*MLL4*	Positive
HCC146	19q13.1	AD000671.1	*MLL4*	Positive
	7p14_15	AC005090.2		
HCC002	19q13.1	AD000671.1	*MLL4*	Positive
HCC003	5p13	AY007685.1	*TERT*	Positive
HCC9907	9q13_21.3	AL133578.1		Negative
HCC155				Positive
H20				Positive
H54	18p11.3	AP000845.4	*NMP p84*[a]	Positive
H120				Positive
HCC143	19q13.1	AD000671.1	*MLL4*	Positive
H49				Positive
H62				Negative
H70				Positive
H72				Positive
H78				Positive
H89				Positive
H76				Positive
H57				Negative
H71				Positive
H85				Positive
H86				Negative
H87				Positive
HCC128				Positive
HCC147				Positive
HCC127				Positive
HCC001				Positive
H148				Negative
H149				Negative
H150				Negative
HCC9833				Negative
HCC9901				Negative
HCC9906				Negative

[a]This integration was already reported. Chromosome locations, GenBank accession numbers, and gene names are indicated for eight viral/cellular junctions from seven HCC samples. Detection of t(17;19)(p11;q13.1) was indicated as positive.

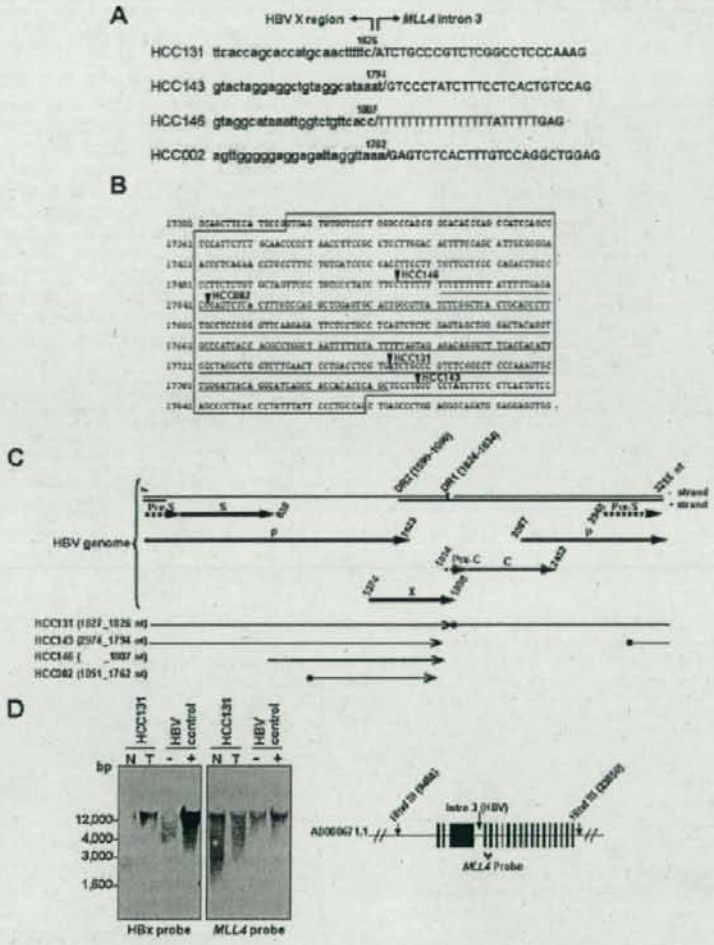
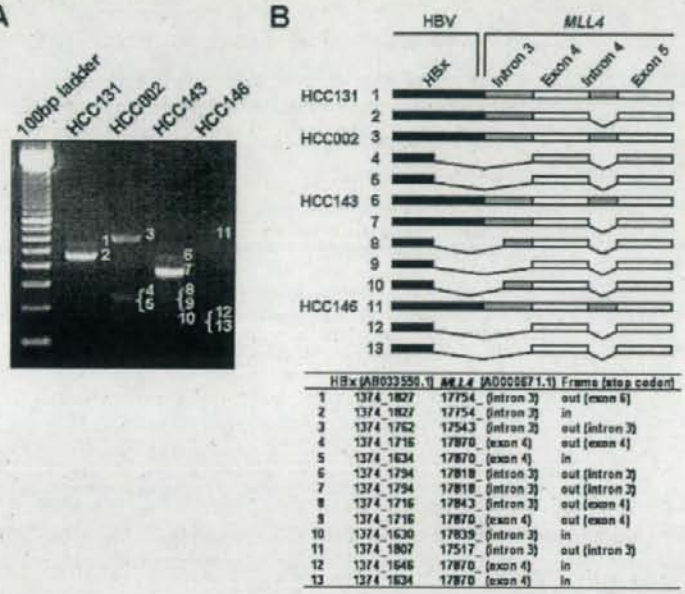


FIGURE 1. **Surrounding sequences of HBV integration sites. A:** Sequences of HBV/cellular DNA junctions in the *MLL4* gene in HCC131, HCC143, HCC146, and HCC002. In each sample, the small letters on the left show sequences of the integrated HBV DNA and capital letters on the right show the flanking *MLL4* gene sequences. Numbers above indicate HBV nucleotides at the HBV/cellular DNA junction (Accession number AB033550.1). **B:** Sequences around intron 3 of the *MLL4* gene and four HBV DNA integration sites are shown. The left side indicates nucleotide positions of *MLL4* gene (accession number AD000671.1). Intron 3 of the *MLL4* gene is indicated by the box (17316_17869 nt). The Alu repeat is shown by underline (17521_17812 nt). **C:** Schematic representation of gene organization of HBV genome and four integrated HBV genomes (HCC002, HCC131, HCC143, and HCC146). Open reading frames and their directions of transcription are represented by an arrow. The numbers above the arrow indicate location of each open reading frame (Accession number AB033550.1). DR1 and DR2 are the 11 basepair direct repeats. • and > indicate the 5′ and 3′ end of integrated HBV DNA sequences (we could not obtain the 5′ end for HCC146). The lengths of the solid lines represent the size and location of the integrated HBV. **D:** Southern blot analysis, using the HBx region as probe (left panel) and the *MLL4* probe (right panel). Hind III-digested DNAs from nontumor tissue of HCC 131 (lane 1), tumor tissue of HCC131 (lane 2), colon cancer tissue as negative control (lane 3), and the HBV integrated HCC tissue as positive control (lane 4). Schematic representation for *MLL4* gene and Hind III site are shown. HBV integration site (intron 3) and *MLL4* probe are indicated. Closed boxes indicate exons of *MLL4* gene.

with an *MLL4* probe of the same membrane, hybridization signals were also detected in the tumor tissue of the patient (Fig. 1D).

HBV Integration Into the *MLL4* Gene Drives Expression of Chimeric Transcripts

In the four HBV/*MLL4* samples, all the integrated viral genome contained HBx promoter and HBx ORF (1374_1838 nucleotides of AB033550.1) except the C-terminus (Fig. 1C). RT-PCR study for detecting fusion transcripts was carried out with HBx primer and reverse primers on various exons of *MLL4* in the four HCC tissues showing various species in each sample (Fig. 2A). In all HCC tissues, in-frame chimeric transcripts that contained exon 4 and exon 5 of *MLL4* were detected (Fig. 2B). In HCC131, two transcripts were observed; one transcript, a major form, showed in-frame fusion containing intron 3 and the other transcript retained intron 4 that led to the creation of the termination codon in exon 6. In HCC002, three transcripts were observed; one transcript



	HBx (AB033550.1)	MLL4 (AB000671.1)	Frame (stop codon)
1	1374_1827	17754_(intron 3)	out (exon 5)
2	1374_1827	17754_(intron 3)	in
3	1374_1762	17543_(intron 3)	out (intron 3)
4	1374_1716	17870_(exon 4)	out (exon 4)
5	1374_1634	17870_(exon 4)	in
6	1374_1794	17818_(intron 3)	out (intron 3)
7	1374_1794	17818_(intron 3)	out (intron 3)
8	1374_1716	17843_(intron 3)	out (exon 4)
9	1374_1716	17870_(exon 4)	out (exon 4)
10	1374_1630	17839_(intron 3)	in
11	1374_1807	17517_(intron 3)	out (intron 3)
12	1374_1646	17870_(exon 4)	in
13	1374_1634	17870_(exon 4)	in

FIGURE 2. RT-PCR analysis of HBx/*MLL4* fusion transcripts. **A:** Various transcripts were observed for each of the HCC tissues by RT-PCR. **B:** Schematic representation of the fusion transcripts from four HCC tissues (HCC131, HCC002, HCC143, and HCC146), and adjacent sequences between HBx (3′ end) and *MLL4* (5′ end) are summarized. HBx cDNA (black boxes) and *MLL4* gene (exon 4 and 5 as white boxes and intron 3 and 4 as gray boxes) are shown. Spliced out sequences are indicated by bars. Location of 5′ end of *MLL4* in intron 3 or exon 4 is also shown. Reading frame based on HBx cDNA followed by *MLL4* is indicated for individual chimeric transcripts. Location of the aberrant stop codon is also shown except for in-frame transcripts. See the Supplementary Appendix for more information.

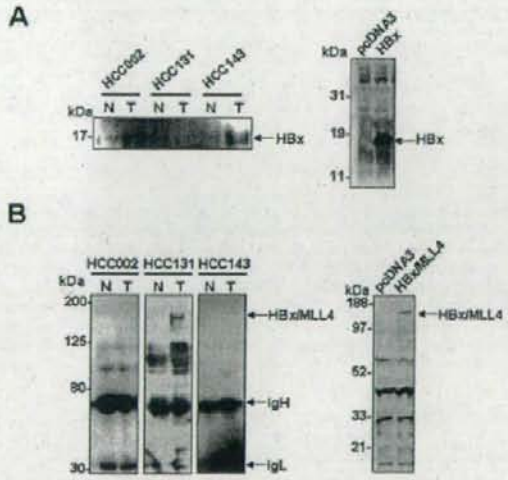


FIGURE 3. **Immunodetection of HBx fusion proteins in HCC samples. A:** Western blot analysis of tumor tissues (T) from HCC002, HCC131, and HCC143 and the adjacent nontumor tissues (N) by using the monoclonal anti-HBx antibody. Recombinant full-length HBx protein expressed in HepG2 cells are also shown. **B:** Immunoprecipitation followed by immunodetection with HBx antibody detected HBx/MLL4 putative fusion protein specifically in HCC131 T (left panel). Western blot using Flag antibody detected an approximately 170-kDa HBx/MLL4 fusion protein transiently expressed in HepG2 cells (right panel). See the Supplementary Appendix for more information.

retained intron 3 causing premature termination and the other two transcripts spliced out intron 3 using distinct 5′ splice sites, resulting in one transcript (nucleotide position 261) showing the in-frame transcript and the other (nucleotide position 343) the premature termination. Similar patterns were observed for HCC146. In HCC143, five species were observed, and the splicing junction of the in-frame transcript was CC-GT, and does not conform to the GT-AG rule.

HBx-MLL4 Fusion Proteins Expressed in HCC

To confirm that the fusion transcripts were translated, the expression of HBx-related proteins in the tumor and adjacent liver tissues were tested by immunodetection with an antibody against HBx protein. Western blot analysis showed that an approximately 17-kDa protein, which represents a short HBx fusion protein compared to recombinant full-length HBx protein expressed in HepG2 cells, is selectively expressed in the HCC002 and HCC143 tumor tissues (Fig. 3A). Immunoprecipitation followed by Western blot analysis detected an approximately 170-kDa protein in HCC131 (Fig. 3B). We constructed an expression vector that can express fusion protein consisting of N-terminally Flag-tagged HBx ORF (amino acids 1_154) and *MLL4* coding region beginning from exon 4 (corresponding to amino acids 820_2715, accession number NM_014727.1), and transiently expressed into HepG2 cells. Western blot using Flag antibody clearly detected an approximately 170-kDa protein (Fig. 3B, right panel). MLL is known to be cleaved at a conserved site and this cleavage generates N- and C-terminal fragments [Hsieh et al., 2003]. MLL4 also possesses a conserved site D/GVDD (amino acids

TABLE 2. cDNA Microarray Results Showing Upregulation and Downregulation by HBx, HBx/*MLL4* Fusion, and Truncated HBx (1_87 aa) Proteins*

	Gene description	Category	HBx		HBx/MLL4		HBx 1_87aa	
			Mean	SD	Mean	SD	Mean	SD
Upregulated gene name								
OR11A1	Olfactory receptor, family 11, subfamily A, member 1	G protein-coupled receptor	3.99	0.02	—		—	
OPN4	Opsin 4 (melanopsin)	G protein-coupled receptor	3.52	0.1	—		—	
UPB1	Ureidopropionase, beta	Hydrolase	3.09	0.52	—		3.89	2.2
HIST1H4L	H4 histone family, member K	Nucleosome structure	3.08	0.3	—		2.72	0.3
HIST1H4I	H4 histone family, member M	Nucleosome structure	2.8	0.02	—		—	
ELL3	Elongation factor RNA polymerase II-like 3	Transcriptional regulation	2.59	0.06	—		—	
BAI1	Brain-specific angiogenesis inhibitor 1	Cell adhesion	2.57	0.62	—		—	
CEP290	Centrosomal protein 290kDa	Centrosomal protein	2.51	0.66	—		—	
HIST1H4B	H4 histone family, member I	Nucleosome structure	2.38	0.01	—		—	
CDC2L1	Cell division cycle 2-like 1	Cell cycle	2.36	0.4	—		2.36	0.04
OR2C1	Olfactory receptor, family 2, subfamily C, member 1	G protein-coupled receptor	2.15	0.15	—		—	
DNCL2B	Dynein, light chain 2B	Motor protein	2.11	0.04	—		—	
ZNF354B	Zinc finger protein 354B	Transcriptional regulation	2.09	0.06	—		—	
MLL4	Mixed-lineage leukemia 4	Transcriptional regulation	—		31	3.25	—	
Downregulated								
AVIL	Advillin	Actin-binding protein	3	0.42	5.2	1.46	3.51	0.68
ENO2	Enolase 2	Hydratase	2.19	0.21	—		—	
KERA	Keratocan	Extracellular matrix	—		4.91	1.37	—	
UBXD1	UBX domain containing 1	Unknown	—		4.76	0.34	—	
PIAS3	Protein inhibitor of activated STAT3	Signal transduction	—		4.68	0.45	—	
MYBPC2	Fast-type myosin binding protein C	Unknown	—		4	0.48	2.48	0.12
PITPNM	Phosphatidylinositol-transfer protein membrane-associated	Cytokinesis	—		3.58	0.83	—	
EHD2	EH-domain containing 2	Endocytosis	—		3.4	0.43	—	
GJB1	Connexin 32	Gap junction	—		3.01	0.62	—	
WASL	Wiskott-Aldrich syndrome-like	Actin polymerization	—		2.48	0.22	—	
TNRC6C	Trinucleotide repeat containing 6c	Unknown	—		2.31	0.28	—	
TBC1D10B	TBC1 domain family, member 10B	Unknown	—		2.19	0.08	—	

*The experiments were performed twice, and the mean and standard deviation (SD) values were determined for each gene.

FIGURE 4. **Reciprocal translocation found in intron 3 of the *MLL4* locus. Chromosomal rearrangement between chromosome 19q13.1 (accession number AD000671.1) and chromosome 17p11.2 (accession number AC087294.18) are shown. The sequences of Alu elements are underlined. The downward pointing arrows on the respective chromosomes indicate the newly synthesized chromosomes (black-to-black and white-to-white) after the recombination events occurred via Alu elements.**

2062_2066), indicating a 170-kDa protein could be a posttranslationally modified product.

Functional Elucidation of HBx/MLL4 Fusion Protein by DNA Microarray

To provide mechanistic insights into molecular etiology such as altered target gene expression regulated by HBx/MLL4 fusion protein, we transiently overexpressed full-length HBx and HBx/MLL4 fusion proteins in HepG2 cells. We employed cDNA microarray technology and identified 13 genes that were upregulated and two genes that were downregulated by HBx protein (Table 2). In contrast, no gene (except for *MLL4* itself) was upregulated and 11 genes were downregulated by HBx/MLL4 fusion protein (Table 2). Uniquely, only one gene, Advillin (*AVIL*) was identified as a common target between HBx and HBx/MLL4 fusion proteins. We checked whether C-terminally truncated HBx protein (amino acids 1_87) could regulate the expression of genes identified by above experiments, because HBx protein in HCC002 and HCC143 only had N-terminal 87 and 86 amino acid residues. Three genes were upregulated and two genes, including *AVIL*, were downregulated by truncated HBx protein (Table 2). Taken together, these data predict that HBx/MLL4 fusion protein would suppress the expression pattern of specific genes.

Alu-Mediated Chromosomal Translocation of *MLL4* to 17p11.2 in HCC

We extended the search for HBV DNA integration into intron 3 of the *MLL4* gene in other HCC samples positive for anti-HBc antibody (Supplementary Table S2). The sequencing analyses failed to detect HBV/*MLL4* DNA sequences, instead demonstrated chimeric sequences between the *MLL4* gene and a particular region of chromosome 17p11.2 (Fig. 4). We detected 22 translocations from 32 HCC samples (Table 1). The sequencing analyses of the translocation products revealed an about 240-bp region at the junction that is highly shared by two chromosomes (approximately 85%) containing Alu elements, suggesting that Alu-mediated homologous recombination facilitated translocation (Fig. 4).

DISCUSSION

The classical mechanism by which tumor-associated viruses contribute to oncogenesis is activation of cellular genes with oncogenic potential through viral genome integration into the cellular genome. HBV genome integration into *SERCA1* (sarco/endoplasmic reticulum calcium ATPase) have been demonstrated [Chami et al., 2000]. The resultant chimeric HBx/SERCA1 protein proposed to be implicated in oncogenesis via an apoptotic

mechanism [Chami et al., 2000]. Reports from two groups, including our observation, demonstrate that the promoter region of the *TERT* gene is targeted by HBV in several HCC tissues [Ferber et al., 2003; Paterlini-Brechot et al., 2003]. Therefore, the *TERT* gene most likely serves as a nonrandom integration site of the viral genome in a subset of HBV-positive HCCs, and the oncogenic HBV DNA integrations may possess the preferential sites. In this study, we further demonstrated four cases of integrations into the *MLL4* gene in HBsAg-positive HCC samples. Sequencing analyses revealed that all of the host sites were within 300 bp of intron 3, flanked with the Alu element of the *MLL4* gene. Recently, HBV DNA integration into *MLL4* gene in three Japanese HCC patients, two cases into exon 3 integration and one into intron 3, were reported [Tamori et al., 2005]. These results support the hypothesis that the oncogenic viral integrations into hepatocytes are not entirely random.

HBV integration into intron 3 of *MLL4* resulted in several fusion transcripts between HBx and *MLL4* that could be directly implicated in liver oncogenesis, albeit the C-terminally truncated HBx protein, as observed in HCC002 and HCC143, might be more closely related to oncogenesis. Our cDNA microarray experiments indicate that HBx/MLL4 fusion protein suppressed the unique genes. It might be speculated that the fusion gene product lacking an AT hook, which is encoded in exons 1–3 of *MLL4*, is directly related to oncogenesis. Further investigation of HBx/MLL4-dependent or N-terminal MLL4-dependent transcriptional regulation may provide a novel insight into the elucidation of etiology of hepatic oncogenesis.

The *MLL4* gene, originally reported as a second human homolog of the *MLL* gene, is mapped to chromosome 19q13.1 [FitzGerald and Diaz, 1999], where gene amplification was reported in HBV-related HCCs [Marchio et al., 1997; Huntsman et al., 1999] and frequent genome rearrangements in solid tumor were reported [Curtis et al., 1998]. We detected the chromosomal translocation of the *MLL4* locus to chromosome 17 in 22 tumors out of 32 samples. The chromosomal rearrangement occurred between intron 3 of the *MLL4* gene of chromosome 19q13.1 and chromosome 17p11.2. The two chromosomal regions share nearly identical Alu elements, indicating that Alu-mediated recombination most likely explains the genome rearrangement. HBV infection and subsequent hepatitis induced DNA damage such as double-strand breaks [Dandri et al., 2002; Bill and Summers, 2004]; therefore, the genome repair mechanism is essential for maintaining the genome integrity and cellular viability.

In conclusion, we detected the translocation breakpoint point in the intron 3 of *MLL4* gene that provides one of the preferential targets for HBV integrations. Indeed we also found recurrent integrations of HBV DNA into intron 3 of *MLL4* gene in four HCC cases, and chimeric HBx/*MLL4* transcripts and HBx/MLL4 proteins, suggesting that the insertional mutagenesis could be functionally relevant to liver oncogenesis.

ACKNOWLEDGMENTS

We are grateful to Dr. Murakami (National Cancer Center Research Institute, Japan) and Dr. Miyamura, Dr. Suzuki, and Dr. Shoji (National Institute of Infectious Disease, Japan) for their technical support and helpful comments. This work was partly supported by grant from the CREST of Japan Science and Technology (II) to I.I.

REFERENCES

Bill CA, Summers J. 2004. Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration. Proc Natl Acad Sci USA 101:11135–11140.

Block TM, Mehta AS, Fimmel CJ, Jordan R. 2003. Molecular viral oncology of hepatocellular carcinoma. Oncogene 22:5093–5107.

Brechot C, Gozuacik D, Murakami Y, Paterlini-Brechot P. 2000. Molecular bases for the development of hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). Semin Cancer Biol 10:211–231.

Chami M, Gozuacik D, Saigo K, Capiod T, Falson P, Lecoeur H, Urashima T, Beckmann J, Gougeon ML, Claret M, le Maire M, Brechot C, Paterlini-Brechot P. 2000. Hepatitis B virus-related insertional mutagenesis implicates SERCA1 gene in the control of apoptosis. Oncogene 19:2877–2886.

Curtis LJ, Li Y, Gerbault-Seureau M, Knick R, Dutrillaux AM, Goubin G, Fawcett J, Cram S, Dutrillaux B, Hanash S, Muleris M. 1998. Amplification of DNA sequences from chromosome 19q13.1 in human pancreatic cell lines. Genomics 53:42–55.

Dandri M, Burda MR, Burkle A, Zuckerman DW, Will H, Rogler CE, Greten H, Petersen J. 2002. Increase in de novo HBV DNA integrations in response to oxidative DNA damage or inhibition of poly(ADP-ribosyl)ation. Hepatology 35:217–223.

Ferber MJ, Montoya DP, Yu C, Aderca I, McGee A, Thorland EC, Nagorney DM, Gostout BS, Burgart LJ, Boix L, Bruix J, McMahon BJ, Cheung TH, Chung TK, Wong YF, Smith DI, Roberts LR. 2003. Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. Oncogene 22:3813–3820.

FitzGerald KT, Diaz MO. 1999. MLL2: a new mammalian member of the trx/MLL family of genes. Genomics 59:187–192.

Gozuacik D, Murakami Y, Saigo K, Chami M, Mugnier C, Lagorce D, Okanoue T, Urashima T, Brechot C, Paterlini-Brechot P. 2001. Identification of human cancer-related genes by naturally occurring hepatitis B virus DNA tagging. Oncogene 20:6233–6240.

Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ. 2003. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol 23:186–194.

Huntsman DG, Chin SF, Muleris M, Batley SJ, Collins VP, Wiedemann LM, Aparicio S, Caldas C. 1999. MLL2, the second human homolog of the Drosophila trithorax gene, maps to 19q13.1 and is amplified in solid tumor cell lines. Oncogene 18:7975–7984.

Marchio A, Meddeb M, Pineau P, Danglot G, Tiollais P, Bernheim A, Dejean A. 1997. Recurrent chromosomal abnormalities in hepatocellular carcinoma detected by comparative genomic hybridization. Genes Chromosomes Cancer 18:59–65.

Mitelman F, Mertens F, Johansson B. 1997. A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. Nat Genet 15:417–474.

Paterlini-Brechot P, Saigo K, Murakami Y, Chami M, Gozuacik D, Mugnier C, Lagorce D, Brechot C. 2003. Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human telomerase gene. Oncogene 22:3911–3916.

Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA. 1995. An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res 23:1087–1088.

Tamori A, Yamanishi Y, Kawashima S, Enomoto M, Tanaka H, Kudo S, Shiomi S, Nishiguchi S. 2005. Alteration of gene expression in human hepatocellular carcinoma with integrated hepatitis B virus DNA. Clin Cancer Res 11:5821–5826.

Available online at www.sciencedirect.com
ScienceDirect
Virology 375 (2008) 424–432
VIROLOGY
www.elsevier.com/locate/yviro

Isolation and gene analysis of interferon α-resistant cell clones of the hepatitis C virus subgenome

Tohru Noguchi [a,b,*], Tomoko Otsubaki [a], Izuru Ando [a], Naoki Ogura [a], Satoru Ikeda [a], Kunitada Shimotohno [b,1]

[a] Central Pharmaceutical Research Institute, Japan Tobacco Inc., Takatsuki, Osaka 569-1125, Japan
[b] Laboratory of Human Tumor Viruses, Department of Viral Oncology, Institute for Virus Research, Kyoto University, Kyoto, Kyoto 606-8507, Japan

Received 5 November 2007; returned to author for revision 27 November 2007; accepted 10 February 2008
Available online 18 March 2008

Abstract

Hepatitis C virus (HCV) proteins appear to play an important role in IFN-resistance, but the molecular mechanism remains unclear. To clarify the mechanism in HCV replicon RNA harboring Huh-7 cells (Huh-9-13), we isolated cellular clones with impaired IFNα-sensitivity. Huh-9-13 was cultured for approximately 2 months in the presence of IFNα, and 4 IFNα-resistant cell clones showing significant resistances were obtained. When total RNA from clones was introduced into Huh-7 cells, the transfected cells also exhibited IFNα-resistance. Although no common mutations were present, mutations in NS3 and NS5A regions were accumulated. Transactivation of IFNα and IFNα-stimulated Stat-1 phosphorylation were reduced, and the elimination of HCV replicon RNA from the clones restored the IFNα signaling. These results suggest that the mutations in the HCV replicon RNA, at least in part, cause an inhibition of IFN signaling and are important for acquisition of IFNα resistance in Huh-9-13.

Keywords: Hepatitis C virus; Replicon; Interferon resistance; Stat-1; Nonstructural protein NS5A

Introduction

Hepatitis C virus (HCV) is the major cause of post-transfusion non-A non-B hepatitis. Approximately 170 million individuals worldwide were estimated to be infected with HCV (Alter, 1997). It has been suggested that the development of liver cirrhosis and hepatocellular carcinoma are consequences of chronic infection with HCV (Hijikata et al., 1993b; Tong et al., 1995).

HCV, a member of the *Flaviviridae* family, has a single-stranded positive-sense linear RNA genome of about 9.5 kb (Hijikata et al., 1991; Kato et al., 1990; Takamizawa et al., 1991). The RNA encodes a single precursor polyprotein of approximately 3010 amino acids (Choo et al., 1991; Okamoto et al., 1991, 1992) that is co- and post-translationally cleaved to produce individual structural (Core, E1, E2) and nonstructural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) by both host and viral proteases (Hijikata et al., 1993a,b; Houghton, 1996).

The cell line Huh-9-13, in which the HCV subgenome can self-replicate, was established by R. Bartenschlager's group (Lohmann et al., 1999). The HCV subgenomic RNA consists of the entire nonstructural coding region of the Con1 strain of the HCV genome, except for the neomycin-resistant gene. This cell line provides significant information for understanding the replication of the HCV genome and is useful as a powerful screening tool for developing anti-HCV drugs (Bartenschlager et al., 2000, 2001).

Interferon alpha (IFNα) is widely used for the treatment of patients with chronic HCV infection; however, the effectiveness of IFNα, especially in genotype 1b, is low at only about 20–30% (Lindsay, 1997), although combination therapy with Ribavirin improves treatment outcomes (up to 50–60%) (McHutchison et al., 1998). According to reports of epidemiologic analysis conducted in Japan, IFN treatment outcomes are related with mutations within a 40 amino acid sequence in NS5A (amino acid

* Corresponding author. Central Pharmaceutical Research Institute, Japan Tobacco Inc., Takatsuki, Osaka 569-1125, Japan. Fax: +81 726 81 9783.
E-mail address: toru.noguchi@ims.jti.co.jp (T. Noguchi).
[1] Current address: Center for Integrated Medical Research, Keio University, Tokyo, Shinjuku 160-8582, Japan.

0042-6822/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.02.009

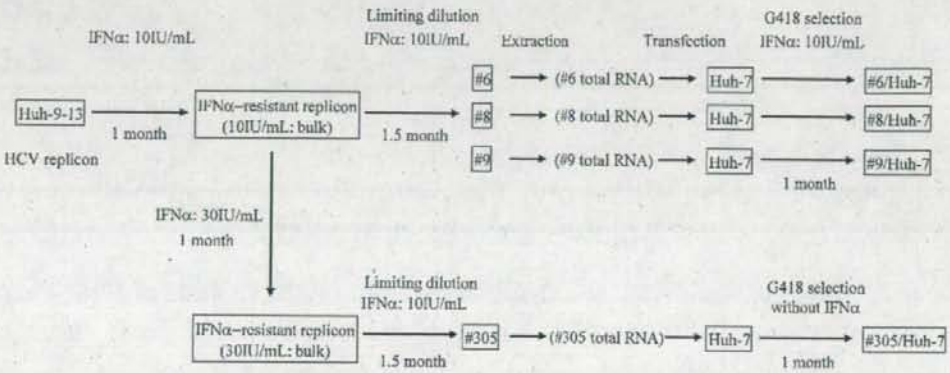


Fig. 1. An outline of the process used for isolation of replicon cells showing IFNα-resistance. Total RNA transfection derived from replicons to naive Huh-7 cells was performed using DMRIE-C reagent (Invitrogen).

(A)

	Cell	EC50(IU/mL)	Fold reduction
Original	Huh-9-13	0.7	1.0
IFNα-resistant	#6	6.9	9.5
	#8	6.7	9.2
	#9	10.2	13.9
	#305	99.2	135.6

(B)

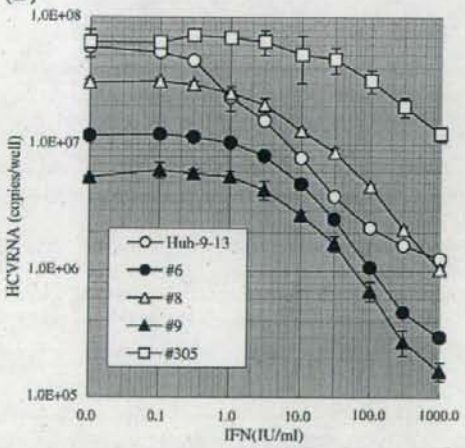


Fig. 2. Reactivity for IFNα in established IFNα-resistant replicon cells (#6, #8, #9, and #305) and original replicon cells (Huh-9-13). The cells were treated with IFNα for 48 h, and the amount of HCV RNA was measured by quantitative RT-PCR. (A) EC_{50} value (IU/mL) of IFNα in each replicon and fold reduction of the value compared to original replicon (Huh-9-13). (B) Change in copy number of HCV RNA in original and IFNα-resistant replicons by IFNα treatment. These experiments were performed in triplicate and mean values are shown.

numbers 2209–2248, based on the sequence of the prototype for HCV-J polyprotein) called the interferon sensitivity determining region (ISDR) (Enomoto et al., 1996). However, it is not clear how NS5A functionally interacts with IFN signals. Alternatively, NS5A is shown to inhibit the activity of double-stranded RNA (dsRNA)-activated protein kinase (PKR) and 2′-5′-oligoadenylate synthetase (2′-5′-OAS) induced by IFNα (Gale et al., 1997; Noguchi et al., 2001; Taguchi et al., 2004).

Recently, Meylan et al. and other groups reported that HCV-NS3-4A protease cleaved Cardif (Meylan et al., 2005) (also designated as VISA (Xu et al., 2005), MAVS (Seth et al., 2005), IPS-1 (Kawai et al., 2005)) and suppressed IFN production through RIG-I signaling. Cardif interacts with RIG-I (Yoneyama et al., 2004) mediated through CARD domains in both molecules

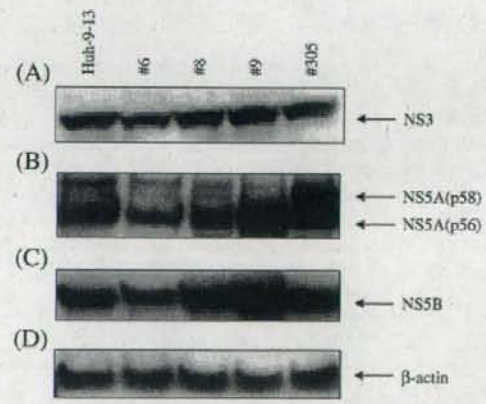


Fig. 3. Western blot analysis of the established IFNα-resistant replicon cells (#6, #8, #9, and #305) and original replicon cells (Huh-9-13). Expression of β-actin was used as an internal control of cellular protein in the replicon cells. Each cell line was inoculated on a 60-mm plate at 3×10^5 cells/well. Twenty-four hours after inoculation, the cells were lysed with SDS sample buffer. Total proteins were subjected to a 2/15% SDS gradient gel, and were subsequently immunoblotted by NS3 (A), NS5A (B), NS5B (C), and β-actin (D) antibody.

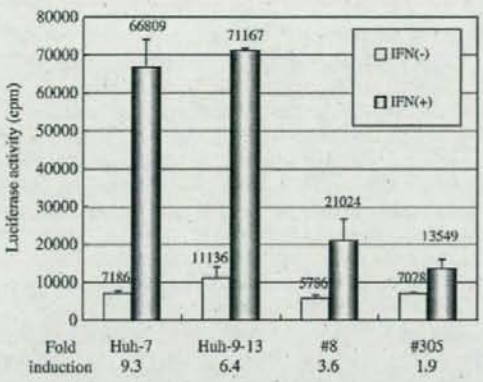


Fig. 4. Transactivation of ISRE in IFNα-resistant replicon cell lines (#8 and #305), original replicon (Huh-9-13), and parental Huh-7 cells by reporter gene (pISRE/Luc) analysis. The cells were stimulated with 1000 IU/mL of IFNα for 24 h after transfection of reporter plasmid DNA. White bars show control (no addition of IFNα) luciferase activity, and black bars show the activity under IFNα stimulation. Values of luciferase activity by IFNα stimulation relative to those of untreated cells are shown below the panel as 'fold induction'.

in a dsRNA-dependent manner, and transduce IFN production signals through the activation of nuclear factor κB (NFκB) and interferon regulatory factor 3 (IRF-3).

Despite bearing an HCV-1b genotype-derived replicon with mutations in ISDR, the replicon cells do not show resistances to IFN (Frese et al., 2002; Guo et al., 2001, 2004). Concerning this point, some reports regarding IFN-resistance acquisition and analysis of this property in the replicon cells (Namba et al., 2004; Sumpter et al., 2004; Zhu et al., 2005) showed involvement of various factors such as viral and/or host gene alterations participating in IFNα-resistance in replicon cells.

Here, we isolated IFNα-resistant clones of the HCV subgenome with accumulated mutations, especially in NS3 and NS5A regions. We observed impairment of phosphorylation of Stat-1 in cells bearing the IFNα-resistant HCV replicon. Our findings suggest that NS5A contributes to the acquisition of IFNα-resistant phenotype in HCV replicon cells.

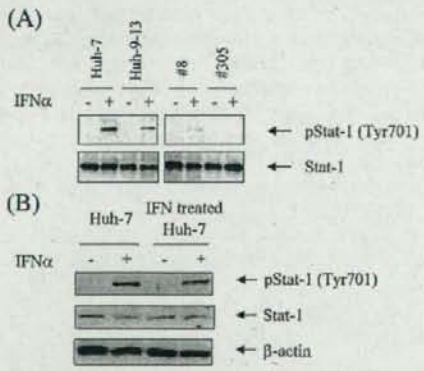


Fig. 5. (A) Change in phosphorylation of Stat-1 in IFNα-resistant replicon cell lines (#8 and #305), original replicon (Huh-9-13) and parental Huh-7 cell. Phosphorylation of Stat-1 was analyzed by western blot analysis using anti-phospho-Stat-1 (Tyr701) antibody. The cells were cultured in medium with or without 500 IU/mL of IFNα for 30 min. Upper panel represents a phospho-Stat-1 (Tyr701) and lower panel shows a Stat-1. Western blot analysis was performed as described in Materials and methods. (B) Change in phosphorylation of Stat-1 in Huh-7 cells maintained in the presence or absence of IFNα (10 IU/mL) for 4 weeks. Upper panel represents a phospho-Stat-1 (Tyr701), middle panel shows a Stat-1 and lower panel shows a β-actin. Phosphorylation of Stat-1 in these cells was examined as described above.

Results

Establishment of IFNα-resistant replicon cell lines

HCV replicon cells were cultured for approximately 1 month in the presence of 10 IU/mL IFNα. HCV RNA titer decreased during the culture; however, the appearance of cells less sensitive to IFNα during prolonged culture was observed by quantitative RT-PCR. The resistant cells were then cloned by limiting dilution. Three clones (Fig. 1: #6, #8, and #9) were obtained, and mixed pools of these resistant cells were further selected in the presence of 30 IU/mL IFNα for another 4 weeks. After confirming decreased sensitivity to IFNα at this dose, the clone

(A)

	Cell	IFN selection	EC50 (IU/mL)	Fold Reduction
Total RNA transfection	Huh-9-13/Huh-7	(-)	0.7	1.0
	#305/Huh-7	(-)	4.1	5.9
original	Huh-9-13	(-)	0.6	1.2

(B)

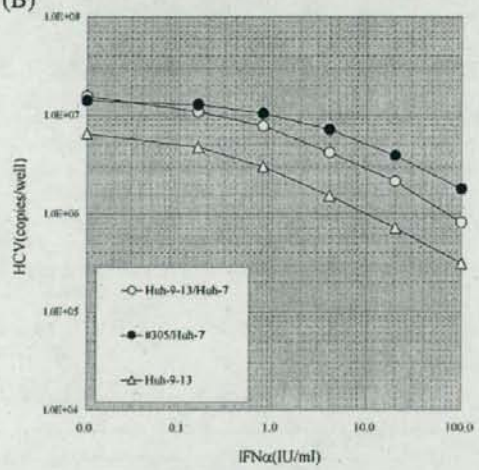


Fig. 6. Reactivity for IFNα in the Huh-7 cells, #305/Huh-7, transfected with total RNA of #305 replicon cells and in the Huh-7 cells, Huh-9-13/Huh-7, transfected with total RNA of original replicon cells (Huh-9-13). These transfected cells were selected with G418 in the absence of IFNα. The amount of HCV RNA was analyzed by quantitative RT-PCR, as described in Fig. 2. (A) EC_{50} value (IU/mL) of IFNα in Huh-9-13/Huh-7 and #305/Huh-7. (B) Change in copy number of HCV RNA in Huh-9-13/Huh-7 and #305/Huh-7 by IFNα treatment. These experiments were performed in triplicate and mean values are shown.