Integrative Transcriptome Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma

Yujin Hoshida,[1,2] Sebastian M.B. Nijman,[1,3] Masahiro Kobayashi,[6] Jennifer A. Chan,[1,7] Jean-Philippe Brunet,[1] Derek Y. Chiang,[1] Augusto Villanueva,[8] Philippa Newell,[10] Kenji Ikeda,[6] Masaji Hashimoto,[6] Goro Watanabe,[6] Stacey Gabriel,[1] Scott L. Friedman,[10] Hiromitsu Kumada,[6] Josep M. Llovet,[8,9,10] and Todd R. Golub[1,2,3,4]

[1]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts; [2]Pediatric Oncology, Dana-Farber Cancer Institute; [3]Children's Hospital Boston, Harvard Medical School; [4]Howard Hughes Medical Institute, Boston, Massachusetts; [5]Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria; [6]Toranomon Hospital, Tokyo, Japan; [7]University of Calgary, Calgary, Alberta, Canada; [8]Barcelona-Clínic Liver Cancer Group, Liver Unit, CIBERehd, Hospital Clínic, IDIBAPS; [9]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain; and [10]Liver Cancer Program, Mount Sinai School of Medicine, New York, New York

Abstract

Hepatocellular carcinoma (HCC) is a highly heterogeneous disease, and prior attempts to develop genomic-based classification for HCC have yielded highly divergent results, indicating difficulty in identifying unified molecular anatomy. We performed a meta-analysis of gene expression profiles in data sets from eight independent patient cohorts across the world. In addition, aiming to establish the real world applicability of a classification system, we profiled 118 formalin-fixed, paraffin-embedded tissues from an additional patient cohort. A total of 603 patients were analyzed, representing the major etiologies of HCC (hepatitis B and C) collected from Western and Eastern countries. We observed three robust HCC subclasses (termed S1, S2, and S3), each correlated with clinical parameters such as tumor size, extent of cellular differentiation, and serum α-fetoprotein levels. An analysis of the components of the signatures indicated that S1 reflected aberrant activation of the WNT signaling pathway, S2 was characterized by proliferation as well as MYC and AKT activation, and S3 was associated with hepatocyte differentiation. Functional studies indicated that the WNT pathway activation signature characteristic of S1 tumors was not simply the result of *β-catenin* mutation but rather was the result of transforming growth factor-β activation, thus representing a new mechanism of WNT pathway activation in HCC. These experiments establish the first consensus classification framework for HCC based on gene expression profiles and highlight the power of integrating multiple data sets to define a robust molecular taxonomy of the disease. [Cancer Res 2009;69(18):7385–92]

Introduction

Hepatocellular carcinoma (HCC) affects approximately half a million patients worldwide and is the most rapidly increasing cause of cancer death in the United States owing to the lack of effective treatment options for advanced disease (1). Numerous lines of clinical and histopathologic evidence suggest that HCC is a heterogeneous disease, but a coherent molecular explanation for this heterogeneity has yet to be reported. Genomic approaches to the classification of HCC therefore hold promise for a molecular taxonomy of the disease.

Mutations in the WNT signaling pathway have been found to be common in HCC, but other DNA level classification approaches have proven challenging. This relates to the enormous complexity of the genomic alterations observed in HCC, likely attributable to the accumulation of chromosomal rearrangements resulting from decades of chronic viral hepatitis and cirrhosis. This complexity makes it difficult to identify the causal genetic events promoting HCC development and progression (2, 3). An alternate approach to HCC classification has been to study tumors at the level of their gene expression profiles. Although several such profiling efforts have been reported (4–11), a cohesive view of expression-based subclasses of HCC has yet to emerge. In part, this is because each of the reported studies analyzed different patient populations (most of them small) on a different microarray platform, with a different primary biological or clinical question in mind. Perhaps not surprisingly then, each study reported a somewhat different view of the heterogeneity of HCC, and it has been therefore impossible to see whether there exists a common biological thread that links these disparate studies.

We believe that any biologically or clinically meaningful classification system should be informative across multiple patient populations and should be independent of any particular microarray platform. In the present study, we therefore set out to define molecular subclasses of HCC that existed across all available HCC data sets, including eight previously reported studies and one new one reported here, totaling 603 patients. We report that indeed there exist three distinct molecular subclasses of HCC that are present in all nine data sets examined, regardless of technical differences between the microarray platforms used to generate the profiles. We show that these subclasses are correlated with histologic, molecular, and clinical features of HCC, and we highlight the important role of transforming growth factor-β (TGF-β) signaling in one of the HCC subclasses. These findings thus create a solid foundation for HCC classification on which to build informed clinical trials for patients with HCC and also suggest new opportunities for therapeutic intervention.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Todd R. Golub, Broad Institute of Massachusetts Institute of Technology and Harvard University, 7 Cambridge Center, Cambridge, MA 02142. Phone: 617-324-0700; Fax: 617-258-0903; E-mail: golub@broad.mit.edu.

doi:10.1158/0008-5472.CAN-09-1089

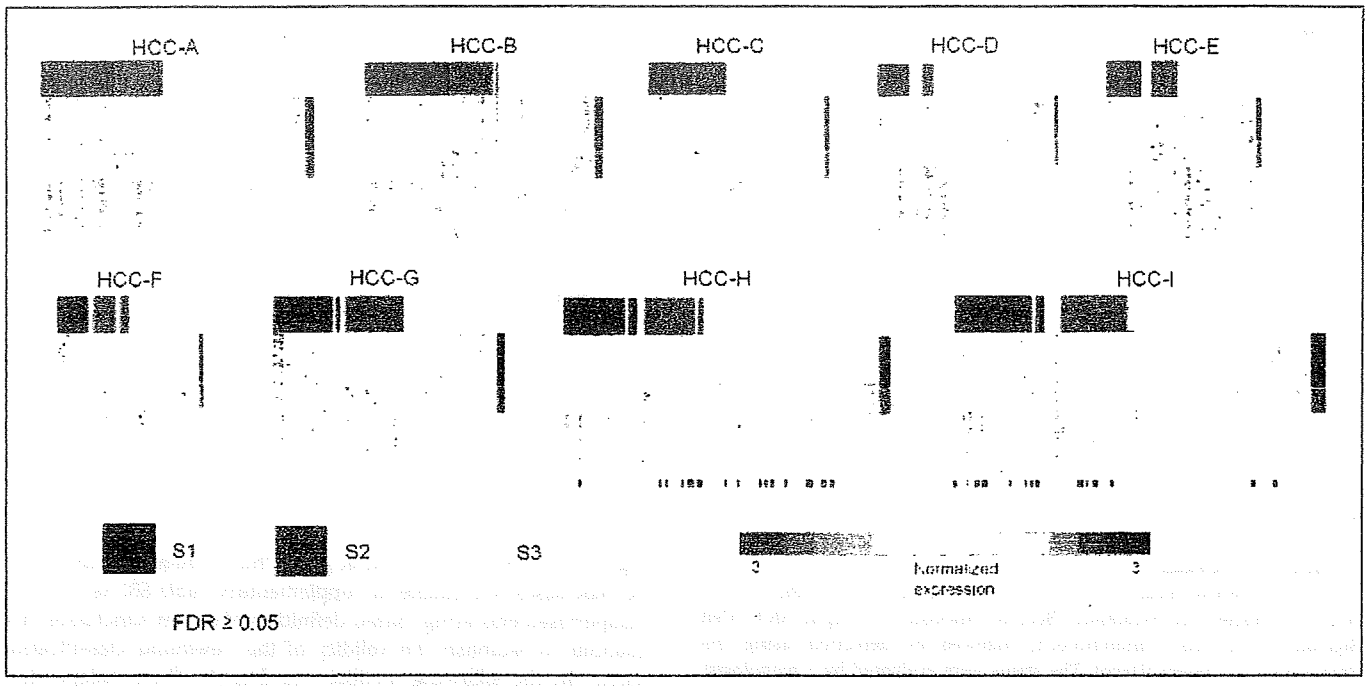

Figure 1. HCC subclasses predicted in nine independent data sets. Predicted subclasses are shown in red (S1), blue (S2), and yellow (S3) with expression pattern of the HCC subclass signature. The proportion of the cases with confident prediction (FDR < 0.05) in HCC-A, HCC-B, HCC-C, HCC-D, HCC-E, HCC-F, HCC-G, HCC-H, and HCC-I were 96%, 96%, 90%, 81%, 79%, 87%, 94%, 83%, and 83%, respectively. Red bars attached to HCC-H and HCC-I indicate positive *β-catenin* mutations and nuclear staining of p53, respectively.

Materials and Methods

Microarray Data Sets and Statistical Analysis

Identification of common HCC subclasses. To define and validate a gene expression–based model of common molecular subclasses of HCC, we collected publicly available gene expression data sets from eight independent cohorts profiled on a wide variety of microarray platforms (HCC-A, HCC-B, HCC-C, HCC-D, HCC-E, HCC-F, HCC-G, and HCC-H; see Supplementary Tables S1 and S2 for details; refs. 4–11). Between the training data sets (HCC-A, HCC-B, and HCC-C) chosen as larger data sets covering major etiologies of HCC to avoid overfitting a model to any particular cohort or microarray platform, corresponding subgroups of the samples were defined by subclass mapping (SubMap) method (12) based on subclasses that identified three unsupervised clustering methods: hierarchical clustering, k-means clustering, and nonnegative matrix factorization, which finds clusters of samples after collapsing the data set into representative "meta-genes" (13).

For each subclass defined by SubMap, meta-analysis marker genes were selected as overexpressed genes compared with the rest of the subclasses (HCC subclass signature) in all the three clustering methods to avoid defining gene expression–based subclasses that were unique to a particular clustering algorithm. Prediction of the subclass was performed for each sample using nearest template prediction method (14, 15) to accommodate the diverse microarray platforms (see Supplementary Data for details).

Molecular annotation of HCC subclasses. Functional characterization of the HCC subclasses was performed using gene set enrichment analysis (GSEA; ref. 16). Two categories of gene sets in Molecular Signature Database (MSigDB)[11] were used: target gene sets regulated by experimental perturbations (377 gene sets) and literature-based manually curated molecular pathway gene sets (150 gene sets).

Clinical data analysis. The hazard of tumor recurrence was calculated to estimate the pattern of HCC recurrence over time after the surgery as previously reported (17, 18). Continuous and proportional data were tested by Wilcoxon rank sum test and Fisher's exact test, respectively. All clinical data analyses were performed using the R statistical package version 2.4.0.[12]

Microarrays for Fixed Tissues

We generated a ninth data set of formalin-fixed, paraffin-embedded (FFPE) tumors (HCC-I), reasoning that any meaningful classification system should be applicable to routinely collected fixed (as opposed to frozen) tissues. We analyzed FFPE tissue blocks from 118 HCC patients who consecutively underwent surgical resection during 1990 to 2001 at Toranomon Hospital. Ethical approval for use of the FFPE tissues, obtained and archived as part of routine clinical care, was acquired from the institutional review board granted on condition that all samples be made anonymous. Total RNA was extracted from macrodissected 10-μm tissue slices (three to four slices for each sample) using High Pure RNA Paraffin kit (Roche). Expression of transcriptionally informative 6,000 genes, selected to capture global state of transcriptome (14), was profiled using DNA-mediated annealing, selection, extension, and ligation (DASL) assay (Illumina; see Supplementary Data; ref. 19).

Microarrays for Cell Lines

Total RNA was isolated using Trizol reagent (Invitrogen) according to the manufacturer's instruction. Microarray experiment was performed using HT_HG-U133A High-Throughput Arrays (Affymetrix). The raw data were normalized using BioConductor affy package.[13]

All microarray data sets are available through National Center for Biotechnology Information Gene Expression Omnibus[14] with accession numbers of GSE10186 (DASL), GSE10393 (U133A), and GPL5474 (transcriptionally informative gene panel for DASL) and our Web site.[15]

[11] http://www.broad.mit.edu/gsea/msigdb/

[12] http://www.r-project.org
[13] http://www.bioconductor.org/
[14] http://www.ncbi.nlm.nih.gov/geo/
[15] http://www.broad.mit.edu/cancer/pub/HCC

Table 1. Clinical phenotypes associated with HCC subclasses

Variable	S1	S2	S3	*P*
Tumor size (cm)*	3.0 [2.0,4.5]	4.5 [2.5,7.0]	2.5 [1.8,4.3]	0.003
Tumor differentiation*				
Well	8 (16%)	4 (10%)	37 (44%)	
Moderate	27 (53%)	23 (59%)	45 (53%)	<0.001
Poor	16 (31%)	12 (31%)	3 (4%)	
AFP (ng/mL)†	50 [14,332]	171 [27,1,251]	13 [5,43]	<0.001
Hepatitis B virus infection‡	39 (38%)	27 (36%)	39 (25%)	0.05
Hepatitis C virus infection‡	55 (53%)	44 (58%)	109 (69%)	0.03

NOTE: Median [25%,75%]. Wilcoxon rank sum test for continuous data, and Fisher's exact test for categorical data.
*HCC-F, HCC-H, and HCC-I: S1, $n = 55$; S2, $n = 46$; S3, $n = 96$.
†HCC-H and HCC-I: S1, $n = 48$; S2, $n = 39$; S3, $n = 83$.
‡HCC-B, HCC-C, HCC-F, HCC-H, and HCC-I: S1, $n = 103$; S2, $n = 76$; S3, $n = 158$.

Immunostaining

Immunohistochemical staining was performed on 10-μm FFPE sections using antibodies for β-catenin (Becton Dickinson), phospho-AKT (Cell Signaling), and p53 (Immunotech) followed by detection using the Envision+ DAB system (Dako). The stains were evaluated by a pathologist blinded to the results of the gene expression profiling, and the results were scored in a binary system. For immunofluorescence staining, cells grown on multiwell chamber slides were fixed by 4% paraformaldehyde and stained for β-catenin (see Supplementary Data for details).

Cell Culture

Human HCC cell lines, Huh-7 (Riken Bioresource Center) and SNU-387 (American Type Culture Collection), were grown in DMEM supplemented with 10% heat-inactivated fetal bovine serum at 37°C in a 5% CO_2 atmosphere. No *β-catenin* mutation has been found in these cell lines.

Western Blotting

Cell lysates were separated on NuPAGE 4% to 12% gels (Invitrogen) and transferred to polyvinylidene difluoride membranes (Bio-Rad) and blotted for α-fetoprotein (AFP: Santa Cruz Biotechnology), β-catenin, phospho-SMAD3 (Cell Signaling), and proliferating cell nuclear antigen (Santa Cruz Biotechnology) antibodies.

β-Catenin Knockdown

Cells were infected with the indicated short hairpin RNA (shRNA) vectors (construct 1, TRCN0000003843; construct 2, TRCN0000003844; The RNAi Consortium[16]) and puromycin selected. Ninety-six hours after infection, cells were counted and seeded in triplicate (20,000/six well). After 10 d, cells were fixed in paraformaldehyde and stained with crystal violet. Dye was extracted with 10% acetic acid and absorbance was determined at 600 nm.

Luciferase Assay

Cells were transfected using Lipofectamine 2000 (Invitrogen) with either TOP-flash or FOP-flash constructs and stimulated with 100 pmol/L TGF-β (R&D Systems) for 48 h. Luciferase activity was measured using Dual-Glo kit (Promega). All transfections were performed in triplicate and measurements were normalized to a SV40 promoter-driven *Renilla* luciferase construct.

Results

Three common molecular subclasses of HCC. The SubMap method identified three robust subclasses in each of the three initial data sets analyzed. We refer to these subclasses as S1, S2, and S3 (Fig. 1), and the complete list of genes that correlated with each of the subclasses is available in Supplementary Table S3. As with any unsupervised clustering–based definition of cancer subclasses, it is essential to establish the validity of the newfound classification system. In the following sections, we describe three independent validations of the three-class structure of HCC. First, we show that the three subclasses are detected with statistical significance in each of the six remaining HCC data sets (totaling 371 patients). Second, we show that the subclasses are associated with clinical parameters. Third, we show that the subclasses are associated with biological mechanism known to be operative in the pathogenesis of HCC.

Statistical validation of subclasses across nine HCC cohorts. As a statistical measure of the validity of the three subclasses, we determined the confidence with which HCC samples could be classified into one of the three subclasses using the HCC subclass signature–based classifier, including 619 genes. As expected, subclass labels were assigned with high confidence [false discovery rate (FDR) < 0.05] to 94% of the training samples (HCC-A, HCC-B, and HCC-C), which were used to define the subclasses in the first place. More importantly, high confidence subclass labels were assigned to 84% of the 371 samples in the validation set (HCC-D, HCC-E, HCC-F, HCC-G, HCC-H, and HCC-I). In contrast, a classifier based on the same number of genes chosen randomly yielded high confidence predictions in <1% of the samples. In addition, our classification system was superior to those reported previously (10, 11, 20, 21) when those classifiers were tested across all of the validation data sets (high confidence predictions using reported signatures were 27–75%; Supplementary Fig. S1). These results, taken together, indicate the statistical significance of our three subclasses and point to the limitation of defining subclasses based on only a single cohort, where overfitting often leads to failure of the classifier to validate on new samples, particularly when profiled on a different microarray platform.

Clinical relevance of HCC subclasses. Having established the statistical validity of the HCC subclasses, we next asked whether any of the subclasses were associated with clinical parameters to add the validity of the subclasses. Clinical data were available for 197 patients, as summarized in Table 1.

Our first observation was that tumors in subclass S2 were larger than the others, whereas tumors in S3 were smaller compared with the rest ($P = 0.003$). In addition, subclass S3 included the majority of well-differentiated tumors (37 of 49; $P < 0.001$), whereas there was no

[16] www.broad.mit.edu/genome_bio/trc/rnai.html

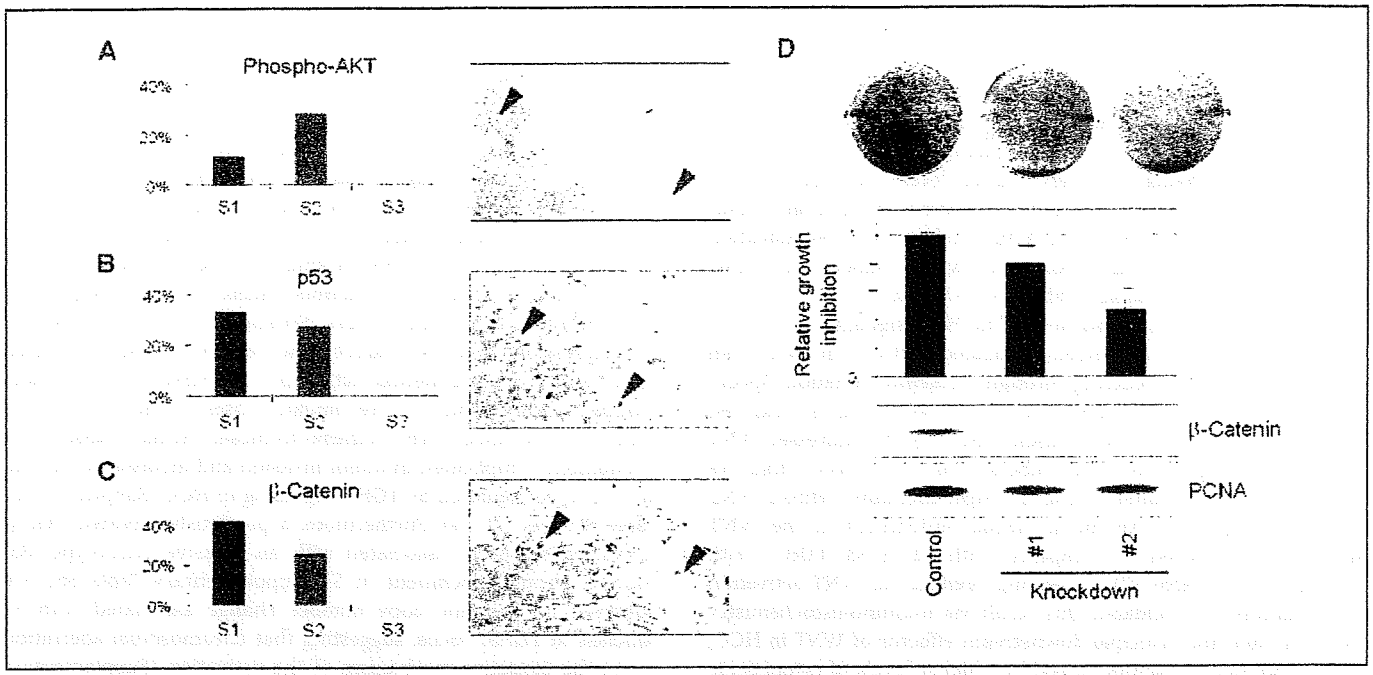

Figure 2. Molecular pathways associated with HCC subclasses. Immunohistochemical analysis of phospho-AKT (*A*), p53 (*B*), and β-catenin (*C*) proteins in HCC-I. *Left*, proportions of the cases with positive staining in each HCC subclass; *right*, representative positive staining (*arrowheads*). Magnification, ×20. *D*, growth inhibition of SNU-387 cells (predicted to be subclass S1) by knocking down β-catenin protein using two different shRNA constructs.

histologic difference between S1 and S2 ($P = 0.73$). We also examined the serum levels of the one clinically used serum biomarker of HCC—AFP. Serum AFP was the highest in S2 (0.001), further supporting the notion that our subclasses are clinically relevant.

Next, we sought to determine whether the HCC subclasses were associated with clinical outcome following surgical resection. We were careful to analyze the two major patterns of HCC recurrence: early recurrence, which is related to residual dissemination of primary tumor cells within the liver, and late recurrence, which is attributable to new primary tumors arising in a hypercarcinogenic state of a cirrhotic liver (17, 18, 22). "Early" recurrence is associated with malignant characteristics of the primary tumor itself, and we reported that it has less effect on patient survival in earlier-stage HCC, in which "late" recurrence is the major determinant of survival (14, 23). We found that subclass S1 was associated with a significantly greater risk of earlier recurrence ($P = 0.03$ within 1 year; Supplementary Fig. S2). This association remained to be significant even after correcting for tumor size in multivariate Cox regression modeling (Supplementary Table S4). Consistent with this observation, S1 tumors exhibited more vascular invasion and satellite lesions (both known risk factors for early recurrence; Supplementary Table S5). These results may suggest that the S1 subclass is associated with a more invasive/disseminative phenotype. Interestingly, we found that a recently reported signature of poor survival defined in patients with more advanced HCC, where early recurrence is the major determinant of survival (4), was associated with S1 and S2, whereas the good survival signature defined in that study was enriched in S3 tumors (Supplementary Table S6), lending further credence to our subclass model. Importantly, our HCC subclasses were not associated with late recurrence, consistent with our recent study indicating that late recurrence is determined not by the characteristics of the tumor itself but rather by the biological state of the surrounding liver at risk (14). Furthermore, consistent with our previous observation, the HCC subclasses showed no association with survival ($P = 0.12$) in our data set (HCC-H) consisting mostly of early-stage tumors.

Molecular pathways associated with HCC subclasses. We next asked whether we could ascribe biological meaning to our validated HCC subclasses. The GSEA results indicated that indeed our HCC subclasses were associated with distinct biological processes, several of which have been implicated in HCC pathogenesis (Supplementary Tables S7 and S8). For example, S2 were tumors associated with a relative suppression of IFN target genes (7), of interest because of the use of IFN as an experimental chemopreventive strategy for HCC. S2 tumors were also enriched in MYC target genes, suggesting that MYC activation is a feature of S2 tumors. Consistent with this observation, we found that a recently reported mouse model of HCC based on MYC overexpression (24) exhibited the S2 subclass signature (Supplementary Fig. S3). S2 tumors were also strongly enriched in a signature of EpCAM positivity (Supplementary Table S6; ref. 25), and in addition, we found that S2 tumors overexpressed *AFP* (consistent with S2 patients having elevated serum AFP levels; Table 1). Lastly, S2 tumors were enriched in a signature of AKT activation (10), and validation experiment indicated a trend toward elevated phosphorylation of AKT as determined by immunohistochemistry ($P = 0.07$; Fig. 2*A*). An AFP-AKT association has been previously observed (10, 26), and we see here that this association is being driven primarily by S2 tumors. The mechanism of AKT activation in these tumors is not known but likely reflects upstream signaling of the phosphatidylinositol 3-kinase (PI3K) pathway (27). As PI3K inhibitors are now entering clinical development, it may be of value to examine their role in S2 tumors in particular.

GSEA also identified differential activation of p53 and p21 target gene sets, with these genes being more abundantly expressed in S3 tumors compared with S1 and S2, consistent with our observation that S3 tumors tend to be lower grade (Table 1). To further validate this result, we performed

immunohistochemistry for p53, wherein nuclear accumulation of p53 protein is well known to reflect inactivating p53 mutation (28). As predicted by the GSEA analysis, S1 and S2 tumors exhibited significantly greater nuclear p53 staining compared with S3 ($P = 0.001$; Fig. 2*B*). The more well-differentiated nature of S3 tumors was also reflected in the S3 gene expression profile, with S3 tumors exhibiting relatively higher levels of expression of hepatocyte function-related genes involved in glycogen/lipid/alcohol metabolism (*APO/ALDH/ADH* family genes), detoxification (*CYP* family genes), coagulation, and oxygen radical scavenging (*CAT, SOD1*; Supplementary Tables S3, S7, and S8).

WNT pathway activation in S1. The WNT signaling pathway is perhaps the best characterized oncogenic pathway in HCC, with pathway activation occurring through *β-catenin* mutation (specifically via mutation in exon 3 in up to 44% of cases) and less frequently in *AXIN1* (<10% of cases: refs. 2, 3). We addressed WNT status with regard to HCC subclass in two ways. First, we performed GSEA analysis using an experimentally defined WNT activation signature. We found strong enrichment of the WNT signature in subclass S1 compared with S2 or S3 (FDR = 0.03; Supplementary Table S7), suggesting preferential WNT activation in S1 tumors. We validated this result via immunohistochemistry for β-catenin (the principal downstream effector of WNT in HCC) and found that S1 tumors indeed had higher levels of cytoplasmic β-catenin protein expression compared with the other subclasses (0.001), again indicating preferential activation of the canonical WNT pathway in S1 (Fig. 2*C*; ref. 29). In addition, we found that shRNAs targeting *β-catenin* resulted in growth inhibition when introduced into the SNU-387 cell line (predicted to be subclass S1), thereby further supporting the hypothesis that WNT activation is functionally important in S1 tumors (Fig. 2*D*).

Mechanisms of WNT activation in S1 tumors. Having determined that S1 tumors exhibit preferential activation of the WNT/β-catenin pathway, we next addressed potential mechanisms for this activity. We first asked whether the S1 tumors were associated with *β-catenin* mutation in HCC-H data set, for which we previously sequenced exon 3 of *β-catenin* gene (11). Surprisingly, *β-catenin* mutations were preferentially found in S3 tumors, consistent with previously reported "CTNNB1" class representing a subset of S3 (11, 30). This result is also consistent with recent evidence indicating that *β-catenin* mutations do not regulate the canonical WNT target genes (e.g., *cyclin D1* and *MYC*) that characterize our S1-associated WNT activation signature (31). These results further suggest that the WNT pathway is activated in S1 tumors by a mechanism other than *β-catenin* mutation.

To explore alternate explanations for WNT pathway activation, we again turned to GSEA, asking whether other gene sets (signatures) enriched in S1 tumors might provide insight into WNT activation in these tumors. Strikingly, we observed strong overexpression of TGF-β target gene sets (i.e., genes expressed as a result of experimental activation of TGF-β) in S1 tumors (Supplementary Table S7). We similarly observed enrichment of a gene set associated with epithelial-to-mesenchymal transition, a phenomenon implicated in tumor invasion and metastasis (32) and known to be regulated by TGF-β signaling in HCC (Supplementary Table S7; refs. 33, 34). Furthermore, a previously reported TGF-β activation signature associated with an invasive phenotype (35) showed strong enrichment in S1 (Supplementary Table S6). We observed no genomic copy number change associated with S1 subclass in *TGFB1* locus, suggesting that chromosomal aberration is not the causative mechanism of the activation (Supplementary Fig. S4). These results indicate that TGF-β and WNT signaling co-occur in the same HCC subclass (subclass S1), and suggest the hypothesis that TGF-β might in some way lead to WNT activity that defines the S1 molecular phenotype.

TGF-β-WNT interactions. We next explored the hypothesis that TGF-β regulates WNT pathway activity in HCC cells. First, we treated the HCC cell line Huh-7 with intact WNT pathway components (classified as subclass S2 and with no activation of S1 and WNT activation signature) with TGF-β and monitored the genome-wide expression consequence. As predicted, TGF-β stimulation induced expression of WNT target genes (FDR < 0.001; Fig. 3*A*) and induced the expression of genes characteristic of subclass S1 (FDR = 0.04; Fig. 3*B*; Supplementary Data) characterized by WNT/

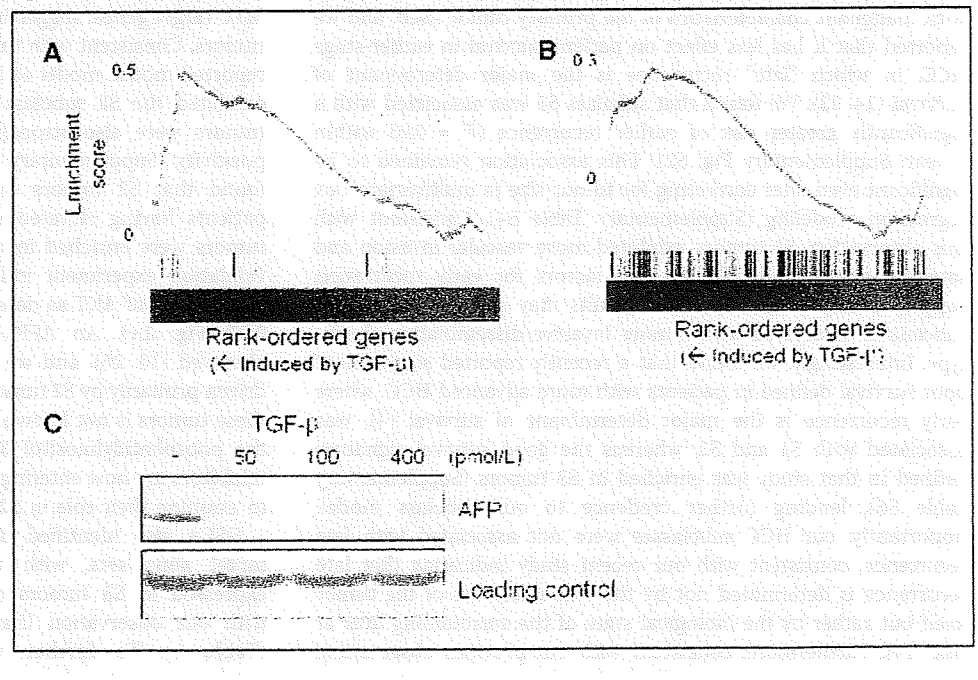

Figure 3. Interaction between WNT pathway and TGF-β. *A*, up-regulation of an experimentally defined WNT target gene set, "KENNY_WNT_UP" (FDR < 0.001), by TGF-β. Genes were rank ordered based on differential expression between TGF-β–treated and untreated Huh-7 cells (predicted to be subclass S2). A database of target gene sets for experimental perturbations (377 gene sets) was assessed by GSEA. *B*, up-regulation of the S1 signature by TGF-β treatment. Genes were rank ordered based on differential expression between treated and untreated Huh-7 cells, and induction of the subclass signature was evaluated by GSEA (FDR = 0.04). *C*, suppression of AFP protein expression by TGF-β treatment. Loading control is nonspecific for AFP antibody to show that equal amounts of protein were loaded.

TGF-β activity while suppressing expression of AFP protein, one of the top markers for S2 (Fig. 3C).

Second, we asked whether TGF-β could regulate the activity of a T-cell factor-lymphoid enhancer factor (TCF-LEF) reporter, further reflecting WNT/β-catenin activity. TGF-β stimulation of Huh-7 cells resulted in activation of a wild-type (TOP-flash) but not mutant (FOP-flash) TCF-LEF luciferase reporter (Fig. 4A). Interestingly, the superactivation of TCF-LEF activity was also observed in the presence of cotransfected mutant β-catenin. These results validate the hypothesis that TGF-β enhances WNT activity in HCC, consistent with the subclass S1 molecular profile.

We next explored the mechanism by which TGF-β augments WNT/β-catenin activity. A simple explanation would be that TGF-β induced expression of β-catenin RNA or protein levels, but we found no evidence for this (Fig. 4B). Strikingly, however, TGF-β treatment resulted in a marked change in β-catenin subcellular localization. Specifically, TGF-β treatment induced a shift from membranous β-catenin staining to a cytoplasmic distribution with focal perinuclear aggregation (Fig. 4C). This suggests that TGF-β enhances WNT signaling by modulating the intracellular pool of free β-catenin.

Taken together, these results validate the observation that TGF-β and WNT activity together typify the S1 subclass of HCC, and further suggest that TGF-β augments WNT activity via alteration of the subcellular localization of β-catenin, consistent with the cross-talk between these pathways observed in other biological contexts (34, 36). This implies that therapeutic cotargeting TGF-β and β-catenin in S1 tumors might be explored as a strategy for the treatment of S1 subclass HCC.

Discussion

Advances in genome technologies are now supporting a breadth of cancer genome characterization studies, including those focusing on HCC. Along with this proliferation of studies has come, however, a certain confusion in the field—different studies often report different results relating to the same set of underlying questions. For example, ~10 articles on the gene expression–based classification of HCC have been published in recent years, but a consensus molecular taxonomy of the disease has yet to emerge. This might lead some to believe that either expression technologies are insufficiently stable or HCC is so hopelessly heterogeneous and complex that regular, reproducible patterns in the data are nonexistent. We report here that in fact a highly reproducible molecular architecture of HCC is identifiable and is detected across all available HCC data sets.

Our analysis of nine HCC data sets totaling 603 patients indicated that there exist three major subclasses of HCC, which we refer to as subclasses S1, S2, and S3 (Fig. 5). Importantly, although the proportion of each subclass varied slightly from study to study, the subclasses were identifiable regardless of the geographic location of the study patients (Asia versus Europe versus United States) or the technology platform used (cDNA versus oligonucleotide arrays, and frozen versus FFPE tissues). Notably, the new data set generated in the present study used FFPE tissues, thereby showing that the three-class structure is readily detectable in specimens collected and stored in the routine clinical setting. This is relevant because the future deployment of diagnostic tests aimed at cancer classification should ideally be applicable to the standard FFPE specimens that are obtained in clinical practice.

Several biological insights can be made from the observed three-class structure of HCC. Class S1 is particularly notable for the prominence of a WNT activation gene expression signature. This is notable because such WNT activation is not simply explained by the presence of activated *β-catenin* mutations, suggesting that additional mechanisms of WNT activation seem to be at play,

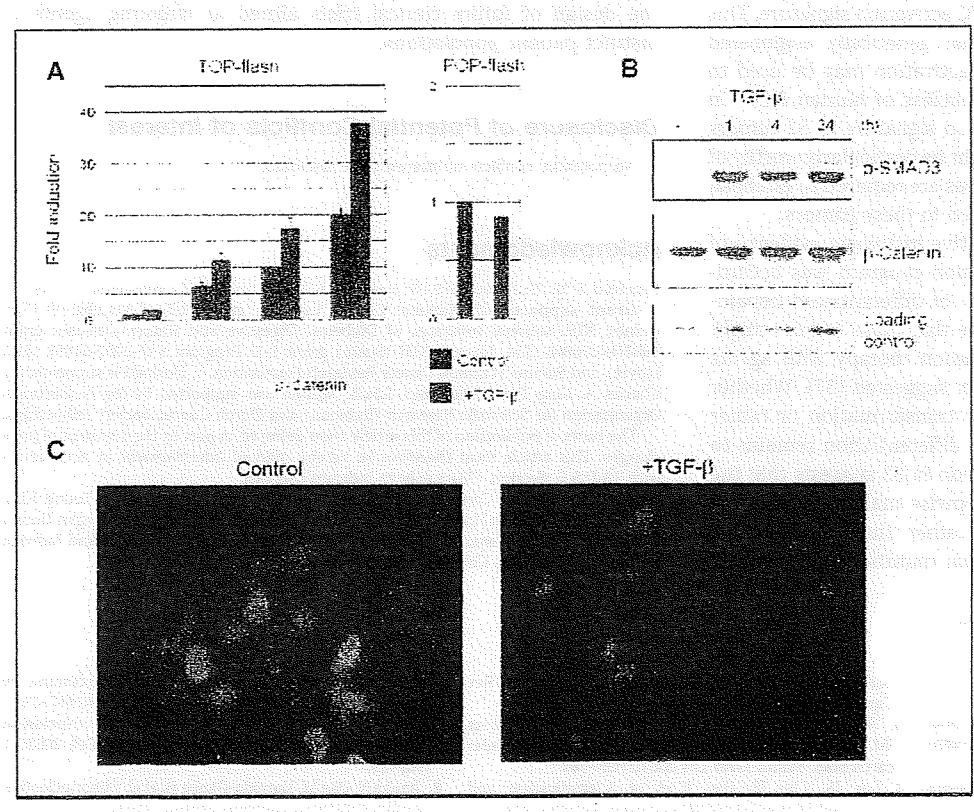

Figure 4. Activation of WNT pathway by TGF-β. *A*, Huh-7 cells were transfected with the indicated reporter constructs and increasing amounts of mutant β-catenin (2, 5, and 10 ng of plasmid). *B*, TGF-β pathway activation was confirmed by phosphorylation of SMAD3. Abundance of β-catenin protein was not changed by TGF-β treatment (100 pmol/L, 48 h). Loading control is nonspecific for phospho-SMAD3 antibody to show that equal amounts of protein were loaded. *C*, Huh-7 cells were stimulated as above and stained for β-catenin. Cellular distribution of β-catenin changed from predominantly membranous to cytoplasmic and perinuclear, and clustered cells spread out with more elongated and flattened morphology.

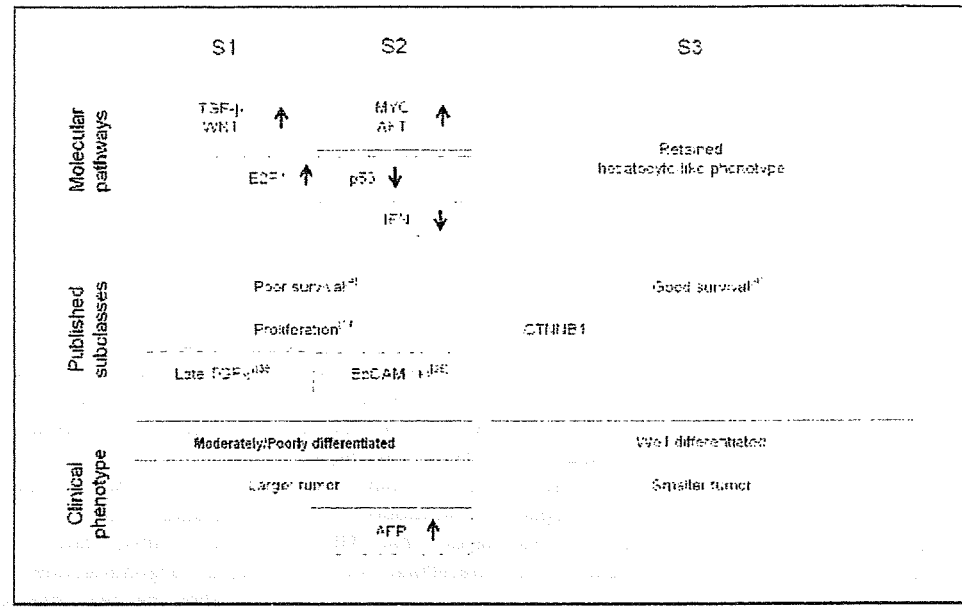

Figure 5. Schematic summary of the characteristics of HCC subclasses.

including TGF-β activation. This may be particularly important in the setting of clinical trials testing β-catenin inhibitors in HCC. Our data suggest that such inhibitors may be worth exploring in HCC beyond those patients harboring *β-catenin* mutation. Although additional mechanistic studies are clearly required, our data support the existence of an interaction between WNT activation and TGF-β activation in S1 tumors, an interaction that has been recently proposed in HCC (34).

Class S2 tumors were notable for their high level of expression of *AFP*, associated with elevated plasma levels of AFP protein compared with non-S2 patients. S2 tumors also tended to be enriched in MYC tumors harboring a MYC activation signature. This is of relevance because it suggests that genetically engineered mouse models of HCC based on MYC activation may be used to interrogate biological basis of the S2 subclass of human HCC. In addition, the finding of an AKT activation signature in S2 tumors suggests that AKT or PI3K inhibitors might be particularly worthy of exploration in this subclass. Further studies are required to establish the mechanism by which AKT is activated in these tumors.

S3 tumors were notable for their relative histologic evidence of differentiation, and the S3 gene expression program was accordingly suggestive of a molecular program of differentiated hepatocyte function. It is tempting to speculate that these tumors might be particularly well suited to differentiation therapy with agents such as retinoids, as has been previously suggested (37). Whether S3 tumors have distinct mechanisms of transformation or rather simply allow for more complete cellular differentiation remains to be determined. The preserved p53 function in S3 suggests that the abrogation of p53 is associated with stepwise malignant transformation of well-differentiated tumors rather than initiation of carcinogenesis. The less frequent *β-catenin* mutations in S1 and S2 may suggest that these tumors arose through different carcinogenic mechanisms compared with S3.

Clearly, much remains to be learned about the biological basis of our observed HCC subclasses. But the fact that they are observed in all studies of HCC examined to date suggests that they represent a reproducible classification framework for the disease. We therefore propose that it will be important to know the subclass of HCC patients entering clinical trials for the treatment of HCC because the response to targeted agents (e.g., β-catenin and PI3K) is likely to be different across the subsets (38). Early observations of differential sensitivity of these distinct tumor types may help guide the design of future clinical trials aimed at targeting agents to distinct patient populations.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 3/25/09; revised 6/9/09; accepted 6/13/09; published OnlineFirst 9/1/09.

Grant support: NIH/National Cancer Institute grant 5U54 CA112962-03 (T.R. Golub), NIH/National Institutes of Diabetes, Digestive and Kidney Diseases grant 1R01DK076986 (J.M. Llovet), NIH (Spain) grant I+D Program SAF-2007-61898 (J.M. Llovet), and Samuel Waxman Cancer Research Foundation. Y. Hoshida is supported by Charles A. King Trust fellowship. S.M.B. Nijman was supported by the Netherlands Organisation for Scientific Research (Rubicon) and Dutch Cancer Society fellowships.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Menno Creyghton for reagents and helpful suggestion; So Young Kim, Ron Firestein, William Hahn, and David Root for the shRNA constructs; Megan Hanna for technical help; Weijia Zhang for critical reading of the manuscript; and Jadwiga Grabarek and Mariko Kobayashi for general support.

References

1. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 2007;132:2557–76.
2. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer 2006;6:674–87.
3. Villanueva A, Newell P, Chiang DY, Friedman SL, Llovet JM. Genomics and signaling pathways in hepatocellular carcinoma. Semin Liver Dis 2007;27:55–76.
4. Lee JS, Chu IS, Heo J, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667–76.
5. Chen X, Cheung ST, So S, et al. Gene expression patterns in human liver cancers. Mol Biol Cell 2002;13:1929–39.
6. Iizuka N, Oka M, Yamada-Okabe H, et al. Oligonucleotide

microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. Lancet 2003;361:923–9.
7. Breuhahn K, Vreden S, Haddad R, et al. Molecular profiling of human hepatocellular carcinoma defines mutually exclusive interferon regulation and insulin-like growth factor II overexpression. Cancer Res 2004;64:6058–64.
8. Ye QH, Qin LX, Forgues M, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nat Med 2003;9:416–23.
9. Midorikawa Y, Tsutsumi S, Nishimura K, et al. Distinct chromosomal bias of gene expression signatures in the progression of hepatocellular carcinoma. Cancer Res 2004;64:7263–70.
10. Boyault S, Rickman DS, de Reynies A, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 2007;45:42–52.
11. Chiang DY, Villanueva A, Hoshida Y, et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res 2008;68:6779–88.
12. Hoshida Y, Brunet JP, Tamayo P, Golub TR, Mesirov JP. Subclass mapping: identifying common subtypes in independent disease data sets. PLoS ONE 2007;2:e1195.
13. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A 2004;101:4164–9.
14. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995–2004.
15. Xu L, Shen SS, Hoshida Y, et al. Gene expression changes in an animal melanoma model correlate with aggressiveness of human melanoma metastases. Mol Cancer Res 2008;6:760–9.
16. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
17. Imamura H, Matsuyama Y, Tanaka E, et al. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. J Hepatol 2003;38:200–7.
18. Mazzaferro V, Romito R, Schiavo M, et al. Prevention of hepatocellular carcinoma recurrence with α-interferon after liver resection in HCV cirrhosis. Hepatology 2006;44:1543–54.
19. Fan JB, Yeakley JM, Bibikova M, et al. A versatile assay for high-throughput gene expression profiling on universal array matrices. Genome Res 2004;14:878–85.
20. Kaposi-Novak P, Lee JS, Gomez-Quiroz L, Coulouarn C, Factor VM, Thorgeirsson SS. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. J Clin Invest 2006;116:1582–95.
21. Lee JS, Heo J, Libbrecht L, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med 2006;12:410–6.
22. Bruix J, Sherman M. Management of hepatocellular carcinoma. Hepatology 2005;42:1208–36.
23. Hoshida Y, Villanueva A, Llovet JM. Molecular profiling to predict hepatocellular carcinoma outcome. Expert Rev Gastroenterol Hepatol 2009;3:101–3.
24. Lee JS, Chu IS, Mikaelyan A, et al. Application of comparative functional genomics to identify best-fit mouse models to study human cancer. Nat Genet 2004;36:1306–11.
25. Yamashita T, Forgues M, Wang W, et al. EpCAM and α-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res 2008;68:1451–61.
26. Sahin F, Kannangai R, Adegbola O, Wang J, Su G, Torbenson M. mTOR and P70 S6 kinase expression in primary liver neoplasms. Clin Cancer Res 2004;10:8421–5.
27. Villanueva A, Chiang DY, Newell P, et al. Pivotal role of mTOR signaling in hepatocellular carcinoma. Gastroenterology 2008;135:1972–83.
28. Vousden KH, Lane DP. p53 in health and disease. Nat Rev Mol Cell Biol 2007;8:275–83.
29. Miller JR, Moon RT. Signal transduction through β-catenin and specification of cell fate during embryogenesis. Genes Dev 1996;10:2527–39.
30. Thorgeirsson SS, Lee JS, Grisham JW. Functional genomics of hepatocellular carcinoma. Hepatology 2006;43:S145–50.
31. Zucman-Rossi J, Benhamouche S, Godard C, et al. Differential effects of inactivated Axin1 and activated β-catenin mutations in human hepatocellular carcinomas. Oncogene 2007;26:774–80.
32. Zavadil J, Bottinger EP. TGF-β and epithelial-to-mesenchymal transitions. Oncogene 2005;24:5764–74.
33. Giannelli G, Bergamini C, Fransvea E, Sgarra C, Antonaci S. Laminin-5 with transforming growth factor-β1 induces epithelial to mesenchymal transition in hepatocellular carcinoma. Gastroenterology 2005;129:1375–83.
34. Fischer AN, Fuchs E, Mikula M, Huber H, Beug H, Mikulits W. PDGF essentially links TGF-β signaling to nuclear β-catenin accumulation in hepatocellular carcinoma progression. Oncogene 2007;26:3395–405.
35. Coulouarn C, Factor VM, Thorgeirsson SS. Transforming growth factor-β gene expression signature in mouse hepatocytes predicts clinical outcome in human cancer. Hepatology 2008;47:2059–67.
36. Jian H, Shen X, Liu I, Semenov M, He X, Wang XF. SMAD3-dependent nuclear translocation of β-catenin is required for TGF-β1-induced proliferation of bone marrow-derived adult human mesenchymal stem cells. Genes Dev 2006;20:666–74.
37. Muto Y, Moriwaki H, Ninomiya M, et al. Prevention of second primary tumors by an acyclic retinoid, polyprenoic acid, in patients with hepatocellular carcinoma. Hepatoma Prevention Study Group. N Engl J Med 1996;334:1561–7.
38. Llovet JM, Bruix J. Molecular targeted therapies in hepatocellular carcinoma. Hepatology 2008;48:1312–27.

Hepatol Int (2009) 3:403–410
DOI 10.1007/s12072-008-9108-8

CASE REPORT

Two cases of development of entecavir resistance during entecavir treatment for nucleoside-naive chronic hepatitis B

Haruhiko Kobashi · Shin-ichi Fujioka · Mitsuhiko Kawaguchi · Hiromitsu Kumada · Osamu Yokosuka · Norio Hayashi · Kazuyuki Suzuki · Takeshi Okanoue · Michio Sata · Hirohito Tsubouchi · Chifumi Sato · Kendo Kiyosawa · Kyuichi Tanikawa · Taku Seriu · Hiroki Ishikawa · Akinobu Takaki · Yoshiaki Iwasaki · Toshiya Osawa · Toshiyuki Takaki · Kosaku Sakaguchi · Yasushi Shiratori · Kazuhide Yamamoto · Daniel J. Tenney · Masao Omata

Received: 27 June 2008 / Accepted: 2 October 2008 / Published online: 9 December 2008

Abstract
Background Entecavir (ETV) is a potent nucleoside analogue against hepatitis B virus (HBV), and emergence of drug resistance is rare in nucleoside-naive patients because development of ETV resistance (ETVr) requires at least three amino acid substitutions in HBV reverse transcriptase. We observed two cases of genotypic ETVr with viral rebound and biochemical breakthrough during

H. Kobashi (✉) · A. Takaki · Y. Iwasaki · K. Sakaguchi · Y. Shiratori · K. Yamamoto
Department of Gastroenterology and Hepatology, Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences, Okayama University, Okayama, Japan
e-mail: hkobashi@md.okayama-u.ac.jp

S.-i. Fujioka · M. Kawaguchi · T. Osawa · T. Takaki
Department of Medicine, Okayama Saiseikai General Hospital, Okayama, Japan

H. Kumada
Center of Liver Disease, Toranomon Hospital, Kanagawa, Japan

O. Yokosuka
Department of Medicine and Clinical Oncology, Graduate School of Medicine, Chiba University, Chiba, Japan

N. Hayashi
Department of Gastroenterology and Hepatology, Graduate School of Medicine, Osaka University, Osaka, Japan

K. Suzuki
First Department of Internal Medicine, Iwate Medical University, Morioka, Japan

T. Okanoue
Molecular Gastroenterology and Hepatology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

M. Sata
Department of Gastroenterology, Kurume University School of Medicine, Fukuoka, Japan

H. Tsubouchi
Digestive Disease and Life-style Related Disease Health Research, Human and Environmental Sciences, Graduate School of Medical and Dental Sciences, Kagoshima University, Kagoshima, Japan

C. Sato
Department of Analytical Health Science, Graduate School of Allied Health Sciences, Tokyo Medical and Dental University, Tokyo, Japan

K. Kiyosawa
Division of Hepatology and Gastroenterology, Department of Internal Medicine, Shinshu University School of Medicine, Matsumoto, Japan

K. Tanikawa
International Institute for Liver Research, Kurume University, Kurume, Japan

T. Seriu · H. Ishikawa
Bristol-Myers Squibb Japan, Pharmaceutical Research Institute, Tokyo, Japan

D. J. Tenney
Bristol-Myers Squibb, Research and Development, Wallingford, CT, USA

M. Omata
Department of Gastroenterology, Faculty of Medicine, University of Tokyo, Tokyo, Japan

Table 1 Baseline characteristics

	Normal range	Unit	Case 1	Case 2
Age	–	–	44 years	47 years
Gender	–	–	Male	Male
T. Bil	0.2–1.0	mg/dl	0.8	0.5
AST	10–40	IU/l	113	48
ALT	5–40	IU/l	199	74
ALP	115–359	IU/l	268	216
BUN	6–20	mg/dl	9.5	15.9
CREA	0.61–1.04	mg/dl	0.95	0.83
ALB	4.0–5.0	g/dl	4.2	4.3
WBC	3500–8500	/μl	6,800	5,650
Hb	13.5–17.0	g/dl	15.7	14.8
PLT	13.1–36.2	10^4/μl	18.9	14.5
Prothrombin time	10–13	second	10.8	11.2
INR		–	1.0	0.9
HBsAg (CLIA)	0–0.05	IU/ml	>100 (positive)	>100 (positive)
anti-HBs (CLIA)	0–10	IU/ml	0 (negative)	0 (negative)
HBeAg (CLIA)	0–1		120 (positive)	190 (positive)
anti-HBe (CLIA)	0–50	%	<35 (negative)	0 (negative)
HBV DNA (PCR)	<2.6	$\log_{10}$ copies/ml	10.0	8.2
HBV genotype			Genotype C	Genotype C
YMDD (sequencing)			YMDD+	YMDD+
			YVDD–	YVDD–
			YIDD–	YIDD–
Liver histology[a]			CH F1/A1	CH F2/A2

[a] Diagnosed according to New Inuyama classification. T. Bil: total bilirubin, AST: aspartate aminotransferase, ALT: alanine aminotransferase, ALP: alkalinephosphatase, BUN: blood urea nitrogen, CREA: serum creatinine, ALB: serum albumin, WBC: white blood cell count, Hb: hemoglobin, PLT: platelet count, INR: international normalized ratio, HBsAg: hepatitis B surface antigen, CLIA: chemiluminescent immunoassay, anti-HBs: antibody to hepatitis B surface antigen, HBeAg: hepatitis B e antigen, anti-HBe: antibody to hepatitis B e antigen, HBV: hepatitis B virus, PCR: polymerase chain reaction, YMDD: tyrosine-methionine-aspartate-aspartate motif, YVDD: tyrosine-valine-aspartate-aspartate motif, YIDD: tyrosine-isoleucine-aspartate-aspartate motif, CH F1/A1: chronic hepatitis with mild fibrosis and mild activity, CH F2/A2: chronic hepatitis with moderate fibrosis and moderate activity

Other baseline characteristics are shown in Table 1. He was enrolled in a phase II clinical trial of ETV and was randomized into 0.1- and 0.5-mg dosage groups. The trial was conducted in Japan in compliance with the ethical principles of the Declaration of Helsinki, Good Clinical Practice guidelines, and Articles/Notifications of the Ministry of Health, Labor and Welfare (H. Kobashi et al., J Gastroenterol Hepatol, in press). He was assigned into the 0.1-mg dosage group and administered ETV at daily dose of 0.1 mg for an initial 52 weeks. Subsequently, he was administered ETV continuously at a daily dose of 0.5 mg for the following 96 weeks. The serum HBV DNA level, which was 10.0 $\log_{10}$ copies/ml at baseline, declined to a nadir of 3.1 $\log_{10}$ copies/ml at week 88 of ETV treatment. Thereafter, HBV DNA level increased from 4.5 $\log_{10}$ copies/ml at week 124 to 6.3 $\log_{10}$ copies/ml at week 140 and 6.7 $\log_{10}$ copies/ml at week 148. ALT levels increased from 28 IU/l at week 144 to 112 IU/l at week 148. The patient discontinued ETV therapy at week 148, and then received a combination therapy of 100 mg of LVD and 10 mg of ADV per day. Afterwards, HBV DNA level dropped to below 2.6 $\log_{10}$ copies/ml and ALT level was normalized after 28 weeks of LVD/ADV dosing (Fig. 1).

HBV DNA sequence analysis was performed using PCR-amplified HBV DNA from preserved serum samples at baseline and at every 24 weeks via HBV DNA polymerase sequence assay (developed at SRL, Inc., Tokyo, Japan). Although sequence analysis of the baseline isolate revealed no substitution in the RT domain of the HBV DNA polymerase gene, analysis of the isolates collected over time revealed the M204I substitution at week 100 and the L180M, S202G, and M204V substitutions at weeks 124 and 144, respectively (Table 2). In addition, a polymorphic residue N238 was found as mixed N238 N/H at week 100 and thereafter. The

ETV treatment of nucleoside-naive patients with chronic hepatitis B (CHB).

Results *Case 1*: A 44-year-old HBeAg-positive man received ETV 0.1 mg/day for 52 weeks and 0.5 mg/day for 96 weeks consecutively. HBV DNA was 10.0 $\log_{10}$ copies/ml at baseline, declined to a nadir of 3.1 at week 100, and rebounded to 4.5 at week 124 and 6.7 at week 148. Alanine aminotransferase (ALT) level increased to 112 IU/l at week 148. Switching to a lamivudine (LVD)/adefovir-dipivoxil combination was effective in decreasing HBV DNA. *Case 2*: A 47-year-old HBeAg-positive man received ETV 0.5 mg/day for 188 weeks. HBV DNA was 8.2 $\log_{10}$ copies/ml at baseline, declined to a nadir of 2.9 at week 124, and then rebounded to 4.7 at week 148 and 6.4 at week 160. ALT level increased to 72 IU/l at week 172. The ETVr-related substitution (S202G), along with LVD-resistance-related substitutions (L180M and M204V), was detected by sequence analysis at week 124 in both case 1 and case 2.

Conclusions ETVr emerged in two Japanese nucleoside-naive CHB patients after prolonged therapy and incomplete suppression and in one patient after <0.5 mg of dosing. ETV patients with detectable HBV DNA or breakthrough after extended therapy should be evaluated for compliance to therapy and potential emergence of resistance.

Keywords Entecavir · HBV · Chronic hepatitis B · Drug resistance · Nucleoside-naive

Introduction

Hepatitis B virus (HBV) infection is a serious health problem because of its high prevalence, estimated to be infecting more than 350 million people worldwide, and its potential for inducing chronic hepatitis, cirrhosis, hepatic decompensation, and hepatocellular carcinoma (HCC) [1, 2]. It has been demonstrated that the most potent risk factor for development of cirrhosis or HCC is serum HBV DNA level [3, 4], and it seems that suppressing serum HBV load is essential for improving the prognosis of HBV carriers. Treatment of chronic hepatitis B (CHB) has evolved markedly with the introduction of nucleoside-analogue antivirals, that is, lamivudine (LVD), adefovir-dipivoxil (ADV), entecavir (ETV), and telbivudine, to clinical practice. LVD, the first approved nucleoside analogue against HBV, was shown to be effective in suppressing HBV DNA replication, improving transaminase levels, improving liver histology, inducing hepatitis B e antigen (HBeAg) seroconversion, and suppressing hepatic insufficiency and hepatocarcinogenesis in CHB and compensated cirrhosis [5, 6]. However, the effectiveness of LVD is limited because of frequent development of drug resistance followed by a hepatitis flare and, occasionally, hepatic failure [7, 8].

ETV, a novel anti-HBV nucleoside analogue, has more than 1,500 times greater potency than LVD in vitro [9]. In clinical trials, ETV administration demonstrated potent anti-HBV activity with a marked decline in serum HBV DNA level and a significant improvement in liver histology than LVD in nucleoside-naive HBeAg-positive and -negative patients [10, 11]. In addition, emergence of ETV resistance (ETVr) or viral rebound was shown in these studies to be rare. From these results, recent treatment guidelines have recommended ETV as the first-line nucleoside analogue for nucleoside-naive CHB patients, including those with cirrhosis [12, 13].

It has been reported that the development of ETVr in nucleoside-naive patients is very rare, even after 4 years of therapy. Recently, however, rare cases of ETVr, which developed in nucleoside-naive patients in clinical studies, have been reported [14–16]. We also observed two patients who developed ETVr-associated HBV reverse transcriptase (RT) substitutions, followed by *virologic rebound*, defined as an elevation in serum HBV DNA of more than 1 $\log_{10}$ copy/ml from nadir, and biochemical breakthrough in long-term ETV treatment of nucleoside-naive CHB patients. In this article, we report these two cases in detail.

Case report

Case 1

A 44-year-old Japanese male CHB patient was positive for hepatitis B surface antigen (HBsAg), HBeAg, serum HBV DNA, and had HBV genotype C, had elevated alanine aminotransferase (ALT) levels, and had no history of nucleoside analogue treatment. The patient had a history of acute appendicitis at age 30, ureteral stone at age 35, and hyperlipidemia at age 43. He had a habit of drinking alcohol (700 ml) daily but did not smoke. At age 27, he was diagnosed for the first time by health screening as an asymptomatic HBV carrier in the immune-tolerant phase, defined by HBsAg positivity and normal liver enzymes, and he was followed up regularly elsewhere with blood tests for liver enzymes. He was found to have ALT elevation. He was referred to our hospital at age 44 and was diagnosed with CHB. Serum HBV DNA level determined by Roche Amplicor™ Monitor PCR assay (lower limit of detection is 2.6 $\log_{10}$ copies/ml = 400 copies/ml; Roche Diagnostics K.K., Tokyo, Japan) [17] was 10.0 $\log_{10}$ copies/ml and serum ALT level was 199 IU/l. Histologic diagnosis by percutaneous liver biopsy at baseline revealed chronic hepatitis with mild fibrosis and mild activity (CH F1/A1, according to the New Inuyama Classification) [18].

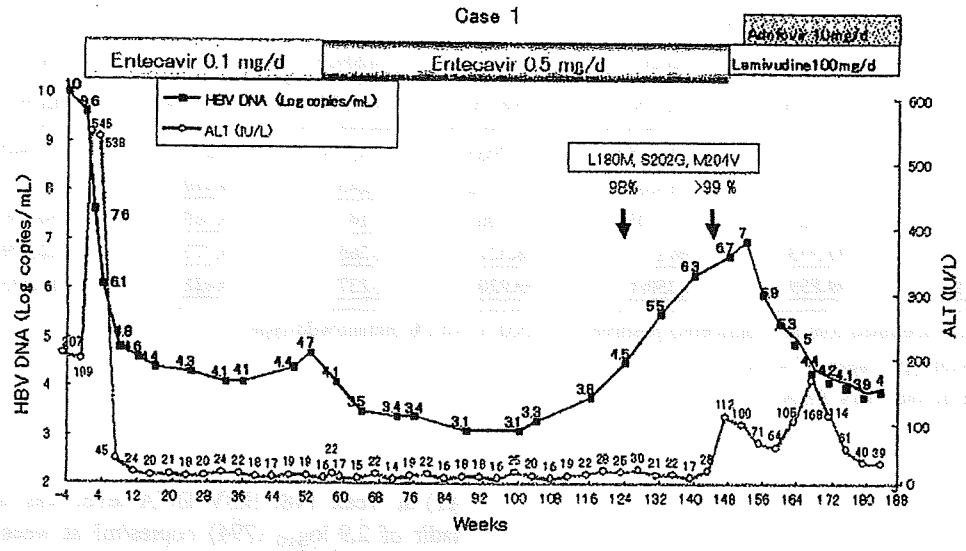

Fig. 1 Clinical course of case 1, a 44-year-old man with nucleoside-naive CHB. ETV treatment reduced ALT levels to below the upper normal limit at week 12 and reduced HBV DNA load to a nadir of 3.1 $\log_{10}$ copies/ml at week 88. However, HBV DNA re-elevated to 4.5 $\log_{10}$ copies/ml at week 124 (virologic breakthrough) and 6.3 $\log_{10}$ copies/ml at week 140, as well as ALT level re-elevated at week 148 (biochemical breakthrough). Sequence analysis of the HBV DNA polymerase gene using serum sample obtained at weeks 124 and 144 revealed the emergence of L180M, M204V (related to LVD resistance), and S202G (related to ETVr) substitutions. SNP-PCR assay revealed that LVDr M204V and ETVr S202G substitutions were detected first at week 124 (98%) and increased at week 148 (>99%). Switching from ETV to LVD/ADV combination treatment at week 148 was successful in reducing HBV DNA load and ALT again

Table 2 Population sequence analysis of isolates from case 1 on ETV therapy

Week	Reverse transcriptase position				
	180	202	204	223	238
0	L	S	M	S/A	N
24	L	S	M	S/A	N
100	L	S	M/I	S/A	N/H
124	M	G	V	S	N/H
144	M	G	V	S	N/H

polymorphic residue S223, which was mixed as S/A at baseline, was found to be only S at weeks 124 and 144.

In addition, preserved serum samples from this patient at baseline and at every 24 weeks were analyzed by an ultrasensitive, single-nucleotide-polymorphism (SNP)-PCR assay, using a method similar to Punia et al. [19] for identification of resistance substitutions, as well as analyzing the sequence of individual clones to determine the genetic linkage of substitutions. SNP-PCR analysis was performed for the two LVD-resistance (LVDr) substitutions, M204V (codon GTG) and M204I (codons ATA and ATT), and the ETVr substitution S202G. Both wild-type and positive control plasmids containing the correct sequence were used at various concentrations to establish the background level as well as the level of detection for each substitution. For clonal analysis, the amplified RT gene from the patient's HBV was cloned into plasmids, as well as 22 to 24 individual clones were selected and sequenced, to determine the genetic linkage of the different substitutions observed.

SNP-PCR analysis for ultrasensitive detection of the resistance substitutions revealed that the LVDr M204V(GTG) and ETVr S202G(GGT) substitutions were not detected (<0.1%) at baseline, week 24, or week 100. The M204I substitution (codon ATA) was detected at low levels at week 24 (0.4%), increased levels at week 100 (6.6%), and was present but at reduced levels at weeks 124 and 148 (0.4% at both time points). The LVDr M204V and ETVr S202G substitutions were detected first at week 124 (98%) and increased levels at week 148 (>99%). The levels of M204I(ATA) were lower at weeks 124 and 144, likely as a result of the dominant M204V/S202G virus (Table 3). Samples at weeks 48 and 76 could not be analyzed conclusively because of low yields of HBV DNA from serum samples.

Clonal analysis revealed that position 223 was a mixture of S and A residues at baseline, the LVDr substitutions L180M and M204V, as well as the ETVr substitution S202G, all emerged simultaneously and were linked in the same virus isolate clones at week 124, isolates that also contained S at position 223. These substitutions did not appear to arise from the LVDr isolates with M204I because the M204I substitution emerged in an isolate with substitution S223A.

Table 3 SNP-PCR analysis of case 1 isolates

Week	M204V		S202G		M204I (ATA)		M204I (ATT)	
	Mut/WT	Ave (%)	Mut/WT	Ave (%)	Mut/WT	Ave (%)	Mut/WT	Ave (%)
0	1/5,424	0.018	1/15,453	0.0065	1/4,199	0.024	1/37,940	0.0026
24	1/5,655	0.018	1/19,000	0.0052	**1/243**	**0.410**	1/46,518	0.0021
100	1/3,846	0.026	1/16,038	0.0062	**1/14**	**6.569**	1/50,456	0.0020
124	**48/1**	**97.973**	**59/1**	**98.327**	**1/265**	**0.377**	1/12,879	0.0078
144	**706/1**	**99.859**	**1,250/1**	**99.920**	**1/237**	**0.421**	1/10,573	0.0095

Cells with bold and underlined font are considered positive (>1/1000 or >0.1% mutant/wild-type)

Mut/WT, mutant/wild type, mean ($N = 3$)

Ave %, average % in total HBV DNA

Case 2

A 47-year-old Japanese male CHB patient was positive for HBsAg, HBeAg, serum HBV DNA, and had HBV genotype C, had elevated ALT levels, and had no history of nucleoside analogue treatment. At age 33, he was diagnosed for the first time as an asymptomatic HBV carrier in the immune-tolerant phase because of positive HBsAg and normal liver enzymes. At age 44, he was found to have ALT elevation, referred to our hospital, and diagnosed with CHB. Histologic diagnosis by percutaneous liver biopsy revealed chronic hepatitis with moderate fibrosis and moderate activity (CH F2/A2 according to the New Inuyama Classification). He was treated with ursodeoxycholic acid at a daily dose of 600 mg orally and glycyrrhizin preparation (stronger Neo-Minophagen C™) 40 ml i.v. thrice per week for 3 months. However, liver enzymes did not normalize. Interferon-α2b administration, three mega units i.m. thrice per week, was started at age 45 and continued for 24 weeks. Although HBV DNA level was reduced transiently to below 3.7 $\log_{10}$ copies/ml at the end of therapy, it rose 9 months after cessation of interferon therapy to 8.2 $\log_{10}$ copies/ml and ALT level increased to 483 IU/l. At age 47, the patient was started on ETV treatment as the subject enrolled in the ETV clinical trial (ETV-053) in Japan at a daily oral dose of 0.5 mg and continued for 188 weeks. A liver biopsy performed 1 month before starting the ETV treatment showed chronic hepatitis with moderate fibrosis and moderate activity (CH F2/A2, according to the New Inuyama Classification). The baseline serum HBV DNA level was 8.2 $\log_{10}$ copies/ml, ALT level was 74 IU/l, and other baseline characteristics were as shown in Table 1. The serum HBV DNA level declined to 3.2 $\log_{10}$ copies/ml and ALT level decreased to below the upper limit of normal at week 32. Liver histology improved to mild-to-moderate fibrosis and mild activity (CH F1-2/A1) at week 48 and chronic hepatitis with mild-to-moderate fibrosis and mild activity (CH F1/A1) at week 148. HBV DNA level was suppressed to a nadir of 2.9 $\log_{10}$ (794) copies/ml at week 124 and rose again to 4.7 $\log_{10}$ copies/ml at week 148, 5.4 $\log_{10}$ copies/ml at week 152, and 6.4 $\log_{10}$ copies/ml at week 160 and 7.0 $\log_{10}$ copies/ml at week 164. ALT level rose to 79 IU/l at week 172 and remained between 40 and 50 IU/l thereafter. ETV at 0.5 mg/day was continued until this time (Fig. 2).

HBV DNA sequence analysis revealed no resistance substitutions in the patient's baseline virus. However, the LVDr-related substitutions L180M and M204V, as well as ETVr-related substitution S202G, were detected at week 124, as a mixed population with wild type, and at week 148, as a pure population (Table 4). In addition, the patient displayed evidence of several polymorphic substitutions at baseline, indicating a mixed quasi-species, which became enriched for those with the resistant virus over time.

SNP-PCR analysis was used to determine the first appearance of the resistance substitutions, using the same method as for case 1. There was no antiviral resistance detected at baseline (<0.1%). The M204V (0.65%) and S202G substitutions were detected first at week 24 but not again until week 124. At weeks 124 and 148, the resistant isolate had become enriched to 43% (M204V) and 98% (M204V), respectively (Table 5).

Clonal analysis was performed to determine the genetic linkage of the various substitutions observed, using the same method as for case 1. The amplified RT gene from the patient's virus was cloned into plasmids, and 24 to 27 individual clones were selected and sequenced. From the clonal analysis, it can be seen that there are three positions that contain mixtures at baseline; position 55 is a mixture of H and R residues, position 221 is a mixture of Y and F residues, and position 269 is a mixture of I and L residues. The substitutions L180M and M204V, as well as the ETVr-related substitution S202G, all emerge simultaneously and in an isolate with H at position 55, Y at position 221, and I at position 269.

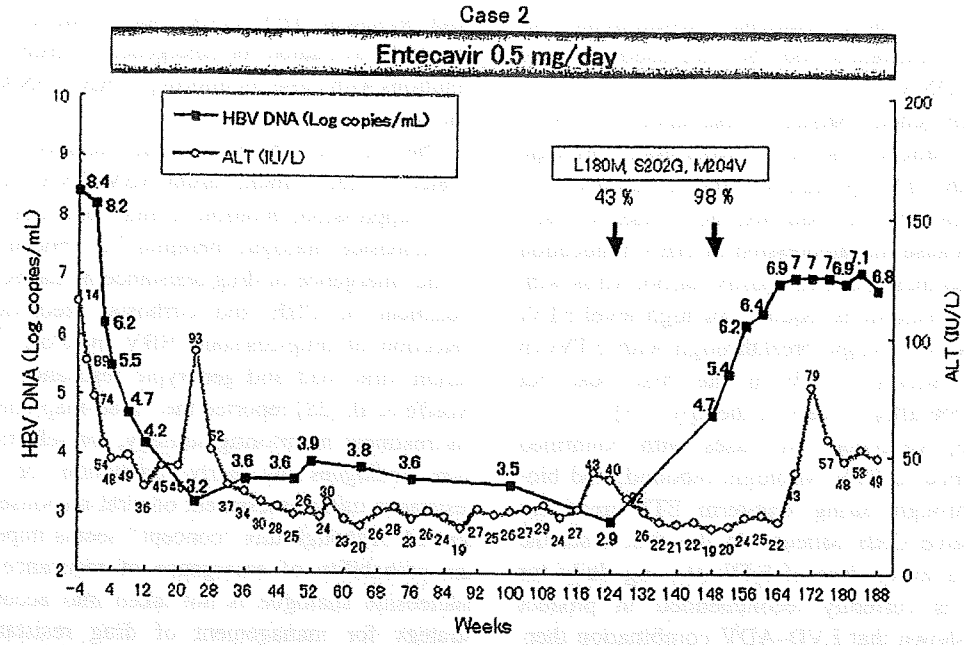

Fig. 2 Clinical course of case 2, a 47-year-old man with nucleoside-naive CHB. ETV treatment reduced ALT level to below the upper normal limit at week 30 and reduced serum HBV DNA level to a nadir of 2.9 $\log_{10}$ copies/ml at week 124. However, HBV DNA level re-elevated to 4.7 $\log_{10}$ copies/ml (virologic breakthrough) at week 148 and 7.0 $\log_{10}$ copies/ml at week 168, as well as ALT level re-elevated to 79 IU/l at week 172. Sequence analysis of the HBV DNA polymerase gene using serum sample obtained at weeks 124 and 148 revealed the emergence of L180M, M204V (related to LVD resistance), and S202G (related to ETVr) substitutions. SNP-PCR assay revealed that the resistant isolate was enriched to 43 (M204V) and 98% (M204V), respectively

Table 4 Population sequence analysis of isolates from case 2 on ETV therapy

Week	RT position								
	55	76	180	191	195	202	204	221	269
0	H/R	S	L	V	F	S	M	Y/F	I/L
24	H/R	S	L	V	F	S	M	Y/F	I/L
52	H/R	S	L	V	F	S	M	Y/F	I/L
100	H	S	L	V	F	S	M	Y/F	I/L
124	H	S/T	L/M	V/I	F/S	S/G	M/V	Y	I/L
148	H	S	M	V	F	G	V	Y	I

Table 5 SNP-PCR analysis of case 2 isolates

Week	S202G[a]	M204V (GTG, %)	M204I (ATA, %)	M204I (ATT, %)
0	Negative	0.016	0.020	0.0065
24	Positive	0.65	0.029	0.018
52	Negative	0.021	0.020	0.018
100	Negative	0.020	0.021	0.010
124	Positive	43	0.33	0.010
148	Positive	98	2.9	0.016

[a] S202G PCR was non-quantitative. A positive indicates 4-fold, 5085-fold, and 10475-fold the wild-type background for weeks 24, 124, and 148, respectively. The baseline isolate gave 1.1-fold the wild-type background

Discussion

The most important limitation of long-term nucleoside analogue treatment for CHB is the emergence of drug-resistant mutant HBV followed by viral breakthrough and hepatitis flare [12]. The most common mutation associated with LVDr involves substitution of methionine in the tyrosine-methionine-aspartate-aspartate (YMDD) motif of the HBV DNA polymerase gene RT domain with valine or isoleucine (M204V/I), with or without a leucine-to-methionine substitution in an upstream region (rtL180M) [20]. It was reported that LVDr was detected at a rate of 14 to 32% after 1 year and 60 to 70% after 5 years of LVD treatment [12]. The substitutions conferring resistance to ADV are asparagine to threonine (N236T) and alanine to valine or threonine (A181V/T) [21], and the cumulative probability of ADV resistance with elevation of HBV DNA level has been reported to be 20% at 5 years in HBeAg-negative patients [22] and as high as 42% in HBeAg-positive patients [23].

In the case of ETV, it has been reported that resistance to the drug requires at least one of three substitutions in HBV RT, that is, rtT184, rtS202, and rtM250, as well as LVDr-related substitutions rtL180M and M204V [24]. Phenotypic analyses of samples associated with virologic breakthrough confirmed that ETV susceptibility correlates

with the spectrum of these additional substitutions conferring genotypic resistance and the increased level of circulating HBV DNA [25].

There is a high genetic barrier to resistance to ETV in nucleoside-naive patients and <1% experience virologic breakthrough with ETVr through 4 years of therapy [15]. However, in LVD-refractory patients, the barrier to resistance is lower because the suppression of HBV replication is not as great and these patients mostly harbor virus with two of the three substitutions required for high-level ETVr [26]. This results in virologic breakthrough with ETVr in LVD-refractory patients at 1% in the first year but increasing to 39.5% after 4 years of therapy [15].

In this article, we report two cases with confirmed genotypic resistance to ETV, virologic rebound, and biochemical breakthrough during long-term ETV treatment for nucleoside-naive CHB patients. In the first case, the patient received a lower dose of ETV (0.1 mg daily for 52 weeks) than is currently recommended in product labeling. It was shown that LVD–ADV combination therapy was apparently effective for the ETV-resistant strain, presumably because there is no cross-resistance between ETV and ADV [26, 27].

SNP-PCR analysis for resistance substitutions revealed that the LVDr M204V(GTG) and the ETVr S202G(GGT) substitutions were negative at baseline and emerged simultaneously at week 124 in both patients. The three resistance substitutions L180M, M204V, and S202G appeared to be genetically linked and did not arise in a stepwise manner in nucleoside-naive patients, as has been described previously.

ETV displays several properties for consideration as the first-line nucleoside analogue because of its potent antiviral activity and a lower frequency of drug resistance than LVD, ADV, or telbivudine [13]. Although ETV is effective in LVD-refractory patients, the potency is reduced somewhat and the barrier to resistance is diminished by the presence of rtM204I/V and rtL180M substitutions. The fact that ETVr may develop in nucleoside-naive patients, even if the chance is small, is noteworthy. In case 1, the patient received a lower dose of ETV (0.1 mg daily), which may be a possible contributing factor to resistance. The common features of our two cases were: HBeAg-positivity, male, high viral load, slow decrease of HBV DNA, and persistently detectable HBV DNA by PCR (>2.6 $\log_{10}$ copies/ml) during the treatment course; however, these characteristics were also present in some other patients who did not develop ETVr. Patient compliance with prescribed therapy also should be assessed in such situations. It is believed that some subpopulations of HBV that proliferate very actively and are not completely suppressed by ETV may have a chance of being selected for the resistance substitutions required for ETV virologic failure. Accordingly, such cases with persistent HBV DNA after extended ETV treatment should be evaluated for emergence of drug-resistance substitutions with close monitoring of HBV DNA level, even in nucleoside-naive patients.

The rate at which resistant mutants are selected is related to pretreatment serum HBV DNA level, rapidity of viral suppression, duration of treatment, and prior exposure to nucleoside analogue therapies [12]. For the management of the emergence of drug resistance in nucleoside analogue treatment of CHB and cirrhosis, prediction and early detection of drug-resistant HBV by close monitoring of serum viral load and genotypic resistance are necessary. Keeffe et al. [28] reported the "road-map concept," that is, on-treatment monitoring strategy, for selection of nucleoside analogues by early prediction of efficacy and resistance using assessment of viral responses at weeks 12 and 24. Although this "concept" seems imperfect because the probability of emergence of resistance to particular nucleoside analogue is not taken into account, a similar strategy for management of drug resistance by close monitoring of viral load and confirming genotypic resistance with consideration of the property of each nucleoside analogue should be established for antiviral treatment of CHB using nucleoside analogues.

Conclusions

We reported two cases of emergence of genotypic resistance to ETV accompanied by virologic breakthrough in nucleoside-naive CHB patients. One patient was treated with a lower than recommended dose of ETV. Although development of ETVr-related gene mutations is rare in nucleoside-naive patients, the patients with a slow decline of HBV DNA or persistent HBV DNA (>2.6 $\log_{10}$ copies/ml) after ETV administration should be evaluated carefully for the potential emergence of ETVr.

Acknowledgments This work was carried out with a grant from Bristol-Myers Squibb. Taku Seriu, Hiroki Ishikawa, and Ronald Rose are employees of Bristol-Myers Squibb. Masao Omata serves as an advisor for Bristol-Myers Squibb. We acknowledge Jie Fang and Kevin Pokornowski, employees of Bristol-Myers Squibb, for genotypic analysis of the patient samples.

References

1. Lavanchy D. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. J Viral Hepat 2004;11:97–107. doi:10.1046/j.1365-2893.2003.00487.x
2. Beasley RP. Hepatitis B virus. The major etiology of hepatocellular carcinoma. Cancer 1988;61:1942–1956. doi:10.1002/1097-0142(19880515)61:10<1942::AID-CNCR2820611003>3.0.CO;2-J

3. Chen CJ, Yang HI, Su J, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. JAMA 2006;295:65–73. doi:10.1001/jama.295.1.65
4. Iloeje UH, Yang HI, Su J, et al. Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. Gastroenterology 2006;130:678–686. doi:10.1053/j.gastro.2005.11.016
5. Lai CL, Chien RN, Leung NW, Chang TT, Guan R, Tai DI, et al. A one year trial of lamivudine for chronic hepatitis B. N Engl J Med 1998;339:61–68. doi:10.1056/NEJM199807093390201
6. Liaw YF, Sung JJ, Chow WC, Farrell G, Lee CZ, Yuen H, et al. Lamivudine for patients with chronic hepatitis B and advanced liver disease. N Engl J Med 2004;351:1521–1531. doi:10.1056/NEJMoa033364
7. Liaw YF, Chien RN, Yeh CT, Tsai SL, Chu CM. Acute exacerbation and hepatitis B virus clearance after emergence of YMDD motif mutation during lamivudine therapy. Hepatology 1999;30:567–572. doi:10.1002/hep.510300221
8. Lok AS, Lai CL, Leung N, Yao GB, Cui ZY, Schiff ER, et al. Long-term safety of lamivudine treatment in patients with chronic hepatitis B. Gastroenterology 2003;125:1714–1722. doi:10.1053/j.gastro.2003.09.033
9. Ono SK, Kato N, Shiratori Y, Kato J, Goto T, Schinazi RF, et al. The polymerase L528M mutation cooperates with nucleotide binding-site mutations, increasing hepatitis B virus replication and drug resistance. J Clin Invest 2001;107:449–455. doi:10.1172/JCI11100
10. Gish RG, Lok AS, Chang TT, de Man RA, Gadano A, Sollano J, et al. Entecavir therapy for up to 96 weeks in patients with HBeAg-positive chronic hepatitis B. Gastroenterology 2007;133:1437–1444. doi:10.1053/j.gastro.2007.08.025
11. Lai CL, Shouval D, Lok AS, Chang TT, Cheinquer H, Goodman Z, et al. Entecavir versus lamivudine for patients with HBeAg-negative chronic hepatitis B. N Engl J Med 2006;354:1011–1020. doi:10.1056/NEJMoa051287
12. Lok AS, McMahon BJ. Chronic hepatitis B. Hepatology 2007;45:507–539. doi:10.1002/hep.21513
13. Keeffe EB, Dieterich DT, Han SH, Jacobson IM, Martin P, Schiff ER, et al. A treatment algorithm for the management of chronic hepatitis B virus infection in the United States: an update. Clin Gastroenterol Hepatol 2006;4:936–962. doi:10.1016/j.cgh.2006.05.016
14. Colonno RJ, Rose R, Baldick CJ, Levine S, Pokornowski K, Yu CF, et al. Entecavir resistance is rare in nucleoside naive patients with hepatitis B. Hepatology 2006;44:1656–1665. doi:10.1002/hep.21422
15. Colonno R, Rose R, Pokornowski K, Baldick C, Eggers B, Yu D, et al. Four year assessment of entecavir resistance in nucleoside naive and lamivudine refractory patients. J Hepatol 2007;46:S294. doi:10.1016/S0168-8278(07)62379-4 (abstract)
16. Suzuki F, Akuta N, Suzuki Y, Yatsuji H, Sezaki H, Arase Y, et al. Selection of a virus strain resistant to entecavir in a nucleoside-naive patient with hepatitis B of genotype H. J Clin Virol 2007;39:149–152. doi:10.1016/j.jcv.2007.03.004
17. Kessler HH, Pierer K, Dragon E, Lackner H, Santner B, Stünzner D, et al. Evaluation of a new assay for HBV DNA quantitation in patients with chronic hepatitis B. Clin Diagn Virol 1998;9:37–43. doi:10.1016/S0928-0197(97)10008-3
18. Ichida F, Tsuji T, Omata M, Ichida T, Inoue K, Kamimura T, et al. New Inuyama classification; new criteria for histological assessment of chronic hepatitis. Int Hepatol Commun 1996;6:112–119. doi:10.1016/S0928-4346(96)00325-8
19. Punia P, Cane P, Teo CG, Saunders N. Quantitation of hepatitis B lamivudine resistant mutants by real-time amplification refractory mutation system PCR. J Hepatol 2004;40:986–992. doi:10.1016/S0168-8278(04)00062-5
20. Lai CL, Dienstag J, Schiff E, Leung NW, Atkins M, Hunt C, et al. Prevalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. Clin Infect Dis 2003;36:687–696. doi:10.1086/368083
21. Angus P, Vaughan R, Xiong S, Yang H, Delaney W, Gibbs C, et al. Resistance to adefovir dipivoxil therapy associated with the selection of a novel mutation in the HBV polymerase. Gastroenterology 2003;125:292–297. doi:10.1016/S0016-5085(03)00939-9
22. Hadziyannis SJ, Tassopoulos NC, Heathcote EJ, Chang TT, Kitis G, Rizzetto M, et al. Adefovir Dipivoxil 438 Study Group. Long-term therapy with adefovir dipivoxil for HBeAg-negative chronic hepatitis B for up to 5 years. Gastroenterology 2006;131:1743–1751. doi:10.1053/j.gastro.2006.09.020
23. Hepsera (Adefovir dipivoxil) current US package insert. CA: Gikead Sciences. p. 3
24. Tenney DJ, Levine SM, Rose RE, Walsh AW, Weinheimer SP, Discotto L, et al. Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to lamivudine. Antimicrob Agents Chemother 2004;48:3498–3507. doi:10.1128/AAC.48.9.3498-3507.2004
25. Baldick CJ, Eggers BJ, Fang J, Levine SM, Pokornowski KA, Rose RE, et al. Hepatitis B virus quasi-species susceptibility to entecavir confirms the relationship between genotypic resistance and patient virologic response. J Hepatol 2008;48:895–902. doi:10.1016/j.jhep.2007.12.024
26. Tenney DJ, Rose RE, Baldick CJ, Levine SM, Pokornowski KA, Walsh AW, et al. Two-year assessment of entecavir resistance in lamivudine-refractory hepatitis B virus patients reveals different clinical outcomes depending on the resistance substitutions present. Antimicrob Agents Chemother 2007;51:902–911. doi:10.1128/AAC.00833-06
27. Villet S, Ollivet A, Pichoud C, Barraud L, Villeneuve JP, Trepo C, et al. Stepwise process for the development of entecavir resistance in a chronic hepatitis B virus infected patient. J Hepatol 2007;46:531–538. doi:10.1016/j.jhep.2006.11.016
28. Keeffe EB, Zeuzem S, Koff RS, Dieterich DT, Esteban-Mur R, Gane EJ, et al. Report of an international workshop: Roadmap for management of patients receiving oral therapy for chronic hepatitis B. Clin Gastroenterol Hepatol 2007;5:890–897. doi:10.1016/j.cgh.2007.05.004

nature genetics

A genome-wide association study identifies variants in the *HLA-DP* locus associated with chronic hepatitis B in Asians

Yoichiro Kamatani[1,2], Sukanya Wattanapokayakit[3], Hidenori Ochi[4,5], Takahisa Kawaguchi[4], Atsushi Takahashi[4], Naoya Hosono[4], Michiaki Kubo[4], Tatsuhiko Tsunoda[4], Naoyuki Kamatani[4], Hiromitsu Kumada[6], Aekkachai Puseenam[7], Thanyachai Sura[7], Yataro Daigo[1,2], Kazuaki Chayama[4,5], Wasun Chantratita[8], Yusuke Nakamura[1,4] & Koichi Matsuda[1]

Chronic hepatitis B is a serious infectious liver disease that often progresses to liver cirrhosis and hepatocellular carcinoma; however, clinical outcomes after viral exposure vary enormously among individuals[1]. Through a two-stage genome-wide association study using 786 Japanese chronic hepatitis B cases and 2,201 controls, we identified a significant association of chronic hepatitis B with 11 SNPs in a region including *HLA-DPA1* and *HLA-DPB1*. We validated these associations by genotyping two SNPs from the region in three additional Japanese and Thai cohorts consisting of 1,300 cases and 2,100 controls (combined $P = 6.34 \times 10^{-39}$ and 2.31×10^{-38}, OR = 0.57 and 0.56, respectively). Subsequent analyses revealed risk haplotypes (*HLA-DPA10202-*DPB1**0501 and *HLA-DPA1**0202-*DPB1**0301, OR = 1.45 and 2.31, respectively) and protective haplotypes (*HLA-DPA1**0103-*DPB1**0402 and *HLA-DPA1**0103-*DPB1**0401, OR = 0.52 and 0.57, respectively). Our findings show that genetic variants in the *HLA-DP* locus are strongly associated with risk of persistent infection with hepatitis B virus.**

Chronic hepatitis B is one of the most common infectious liver diseases caused by hepatitis B virus (HBV). HBV infection shows a marked regional diversity and is very prevalent in the Asia-Pacific region; HBsAg seropositivity rates are as high as 5–12% in Thai and China, but as low as 0.2–0.5% in North America and Europe[2]. It is estimated that, at present, more than 400 million people worldwide are chronically infected with HBV, and nearly 60% of liver cancers are considered to be related to chronic hepatitis B and subsequent liver cirrhosis[3]. Most HBV carriers are considered to have been infected through maternal transmission in the neonatal period or infancy, particularly in Japan[4]. Although some HBV carriers spontaneously eliminate the virus, 2–10% of individuals with chronic hepatitis B are estimated to develop liver cirrhosis every year, and a subset of these individuals suffer from liver failure or hepatocellular carcinoma[1]. Because clinical outcomes after exposure to HBV are highly variable, identification of genetic and environmental factors that are related to progression of HBV-induced liver diseases is critical.

Several epidemiological factors such as age at infection, sex, chronic alcohol abuse[5] and co-infection with other hepatitis viruses[6] were suspected to affect viral persistence. In addition, a twin study in Taiwan indicated that host genetic background influences infection outcome[7]. Although genetic variants in *IFNG*, *TNF*, *VDR*, *ESR1* and several *HLA* loci were shown to associate with chronic hepatitis B[8–12], none of the associations has been proven to be conclusive. To identify disease-predisposing variants, we carried out a two-stage association study for chronic hepatitis B using genome-wide SNPs as genetic markers.

Characteristics of each cohort group are shown in **Supplementary Table 1** online. We carried out a two-stage genome-wide association approach as described in the Methods. In the first stage, we genotyped 179 Japanese individuals with chronic hepatitis B and 934 control individuals using Illumina HumanHap550 BeadChip (**Fig. 1a**). For the second stage, we selected the top 12,000 SNPs that had the smallest *P* values on the basis of minimum *P* value considering three genetic models: allelic, dominant or recessive. Analysis of an independent set of 607 cases and 1,267 controls using these sub-selected SNPs showed 11 SNPs to be significantly associated ($P = 3.62 \times 10^{-8} \sim 1.16 \times 10^{-13}$) with chronic hepatitis B after Bonferroni correction (**Fig. 1b** and **Supplementary Table 2** online). Application of the Cochrane-Armitage

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, the University of Tokyo, Tokyo, Japan. [2]Department of Medical Genome Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Tokyo, Japan. [3]Medical Genetic Section, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand. [4]Center for Genomic Medicine, RIKEN, Kanagawa, Japan. [5]Department of Medicine and Molecular Science, Division of Frontier Medical Science, Programs for Biomedical Research, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima, Japan. [6]Department of Hepatology, Toranomon Hospital, Tokyo, Japan. [7]Department of Medicine, Faculty of Medicine and [8]Virology and Molecular Microbiology Unit, Department of Pathology, Faculty of Medicine, Ramathidi Hospital, Mahidol University, Bangkok, Thailand. Correspondence should be addressed to Y.N. (yusuke@ims.u-tokyo.ac.jp).

Received 10 November 2008; accepted 16 January 2009; published online 6 April 2009; doi:10.1038/ng.348

test to all the tested SNPs indicated that the genetic inflation factor lambda was 1.02 for the second stage (**Supplementary Fig. 1a** online), implying a low possibility of false positive associations due to population stratification. All 11 SNPs are located within or around the *HLA-DPA1* and *HLA-DPB1* locus (**Fig. 2**). We also conducted age- and sex-adjusted analysis using a logistic regression model, and confirmed similar association after adjustment (data not shown).

To validate the result of the discovery-phase analysis, we carried out replication analyses using three independent cohorts. We selected the most or second-most strongly associated SNPs from each *HLA-DP* locus (rs9277535 on *HLA-DPB1* and rs3077 on *HLA-DPA1*, respectively), as we failed to design a Taqman or Invader probe for rs2395309 on *HLA-DPA1*. We first examined two independent sets of Japanese case–control samples comprising 274 cases and 274 controls (age-, sex- and alcohol consumption–matched cohort from BioBank Japan) as well as 718 cases and 1,280 controls. We found significant associations at two SNP loci in both studies ($P = 1.06 \times 10^{-16} \sim 1.96 \times 10^{-6}$; **Table 1**). We also genotyped 308 individuals with chronic hepatitis B and 546 healthy controls in Thailand, and further confirmed the association at the two loci, rs3077 ($P = 6.53 \times 10^{-6}$) and rs9277535 ($P = 6.52 \times 10^{-8}$).

To combine these studies, we conducted a meta-analysis with a fixed-effects model using the Mantel-Haenszel method. As shown in **Table 1** and **Supplementary Figure 1b**, the odds ratios (OR) were quite similar across the four studies (the second stage of GWAS and three replication studies) and no heterogeneity was observed. Mantel-Haenszel *P* values for independence were 2.31×10^{-38} for rs3077 (OR = 0.56, 95% confidence interval (CI) = 0.51–0.61), and 6.34×10^{-39} for rs9277535 (OR = 0.57, 95% CI = 0.52–0.62).

The 11 SNPs showing significant associations are located within a 50-kb region including *HLA-DPA1* and *HLA-DPB1* (**Fig. 2**). Although the *HLA* region is known to show extensive linkage disequilibrium (LD) spanning over 7 Mb, the LD block including these 11 SNPs (surrounded by a bold line in **Fig. 2a**) was not in strong LD with the other *HLA* loci. In accordance with the extent of LD, only SNPs around the *HLA-DPA1* and *HLA-DPB1* genes showed very strong associations with chronic HBV (surrounded by a bold line in **Fig. 2b**), and SNPs outside of this particular LD block did not have significant association.

HLA-DPA1 and *HLA-DPB1* encode the HLA-DP α and β chains, respectively. HLA-DPs belong to the HLA class II molecules that form heterodimers on the cell surface and present antigens to CD4-positive T lymphocytes. HLA-DPs are highly polymorphic, especially in exon 2, which encodes antigen-binding sites. We thus considered that the association of these SNPs with chronic HBV might reflect variations in antigen-binding sites that might affect the immune response to HBV. We genotyped *HLA-DPA1* and *HLA-DPB1* alleles by direct sequencing of exon 2 (cases at second stage and controls at first stage) and found significant association of chronic hepatitis B with *HLA-DPA1**0103, *DPA1**0202, *DPB1**0402 and *DPB1**0501 ($P = 2.93 \times 10^{-11}$, 4.45×10^{-8}, 2.27×10^{-7} and 6.98×10^{-7}, respectively; **Supplementary Table 3** online). Because sequence variants in exon 2 of *HLA-DPA1* and *HLA-DPB1* could be linked to individual nucleotide variants, we inferred haplotypes using the 11 SNPs and variants in exon 2, and found very strong LD among them (**Supplementary Fig. 2**

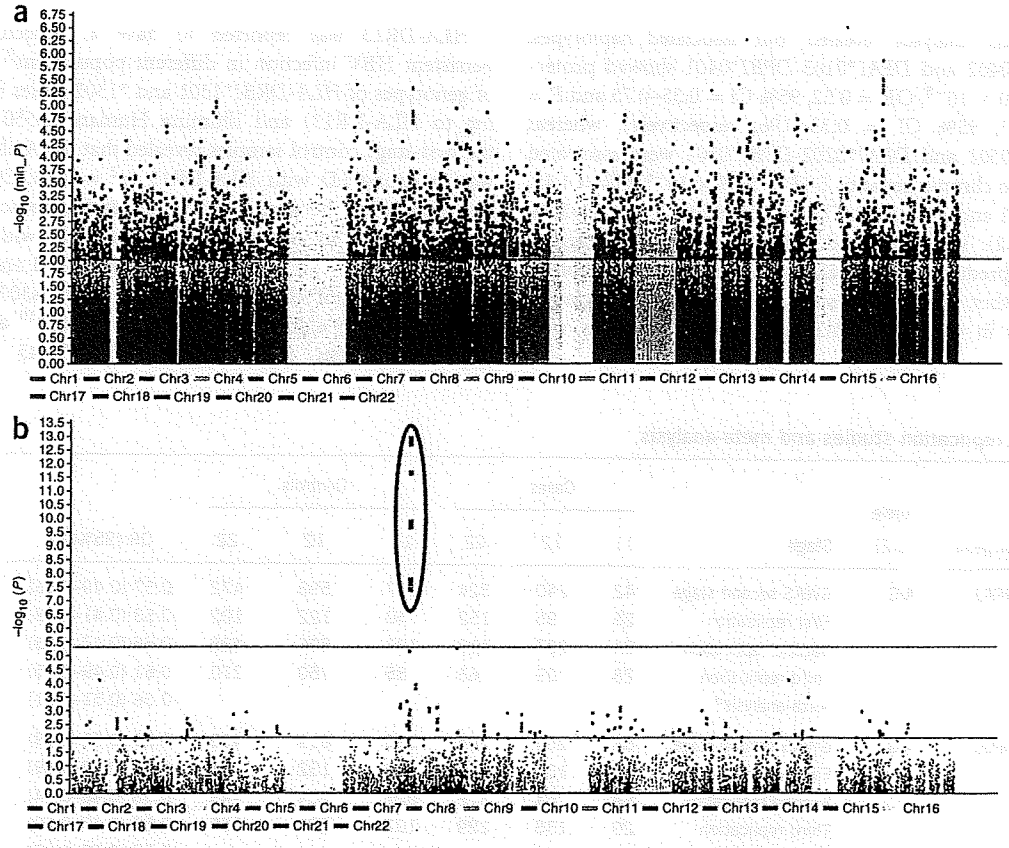

Figure 1 Results from a two-stage genome-wide association study. (**a**) $-\log_{10} P$ value plot at the first stage. Each *P* value is the minimum of Fisher's exact tests for three models: dominant, recessive and allele frequency model. (**b**) $-\log_{10} P$ value plot at the second stage. *P* values were calculated by 1-d.f. Cochrane-Armitage trend test. The large dots circled by red on the chromosome 6 showed significant associations ($P < 5.06 \times 10^{-6}$) with chronic hepatitis B.

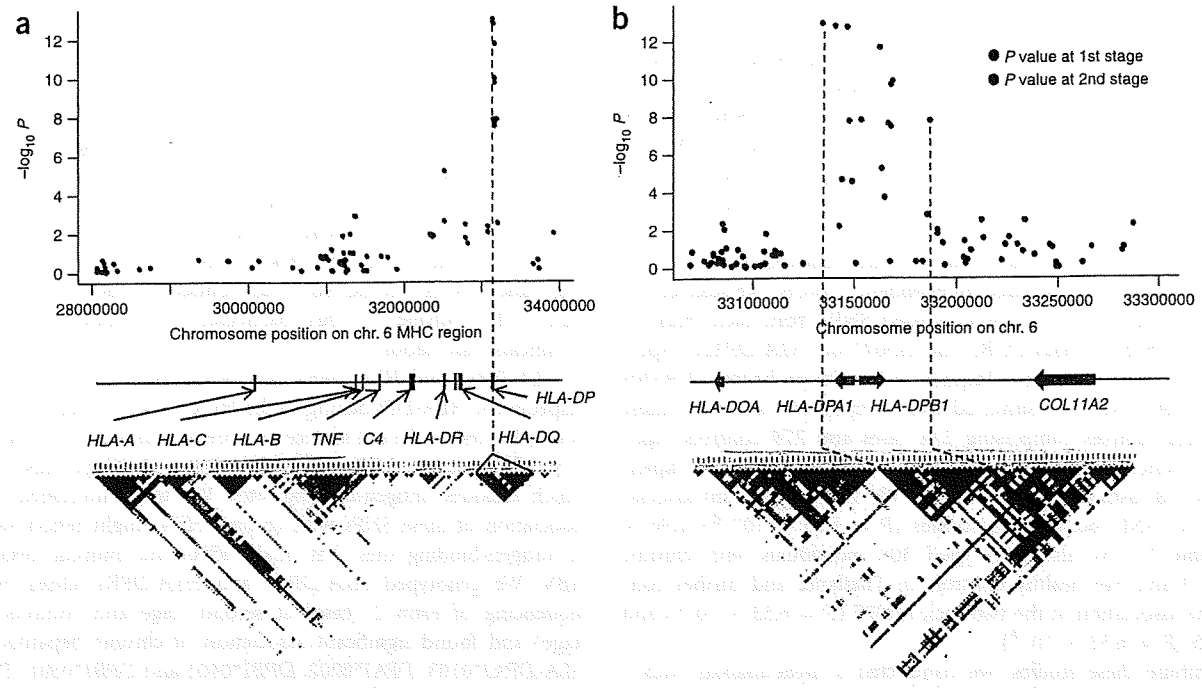

Figure 2 Case–control association results and linkage disequilibrium map of the MHC region. (**a**) *P*-value plot, genomic structure and LD map of the second stage within the extended MHC region of chromosome 6. The LD map based on *D′* was drawn using the genotype data of the cases and the controls in the second stage. (**b**) *P*-value plot, genomic structure and LD map around the *HLA-DPA1* and *HLA-DPB1* region. Black dots and red dots represent *P* values in the first and the second stage, respectively. The LD map based on *D′* was drawn using the genotype data of the cases and the controls in the first stage.

online). Case–control analyses revealed four associated haplotypes: *DPA1**0103-*DPB1**0402 and *DPA1**0103-*DPB1**0401 showed protective effects ($P = 6.00 \times 10^{-8}$, OR = 0.52, 95% CI = 0.35–0.75 and $P = 0.002$, OR = 0.57, 95% CI = 0.33–0.96, respectively), whereas *DPA1**0202-*DPB1**0501 and *DPA1**0202-*DPB1**0301 were associated with susceptibility to chronic hepatitis B ($P = 5.79 \times 10^{-6}$, OR = 1.45, 95% CI = 1.16–1.81 and $P = 0.002$, OR = 2.31, 95% CI = 1.39–3.84, respectively; **Table 2**). We also found various sets of SNPs (tagging SNPs) that could predict *HLA-DP* alleles (**Supplementary Table 4** online). Taken together, our findings strongly implicate an association of genetic variants in the *HLA-DPA1* and *HLA-DPB1* genes with chronic hepatitis B.

HLA-DR13 was reported to have a protective effect against persistent HBV infection in different populations[9,13,14]. Comparison of genotypes of *HLA-DRB1**1301 and *1302 alleles (both corresponding to HLA-DR13) and Illumina HumanHap550 SNPs in 333 of the first-stage control samples revealed that the A allele of rs11752643 was in strong LD with HLA-DR13 ($r^2 = 0.83$, $D' = 1$). However, the association between rs11752643 and chronic hepatitis B was not significant in our second stage GWAS, with an uncorrected *P* value of 1.04×10^{-4} (**Supplementary Table 5** online). In addition, the association of chronic hepatitis B with rs3077 and rs9277535 remained highly significant ($P = 2.11 \times 10^{-10}$ and 1.73×10^{-9}, respectively) after adjustment for rs11752643 using a logistic

Table 1 Results of replication studies and meta-analysis

SNP	Nearest gene	Allele (1/2)	Stage	Cases			Controls			OR (95%CI)[a]	P[b]	P_{het}[c]
				11	12	22	11	12	22			
rs3077	*HLA-DPA1*	A/G	GWAS second stage	42	240	324	197	598	472	0.57 (0.49–0.66)	1.26E–13	
			First replication	25	95	152	50	122	102	0.53 (0.41–0.69)	1.73E–06	
			Second replication	64	237	410	197	596	485	0.55 (0.47–0.63)	1.06E–16	
			Third replication	28	109	163	85	250	210	0.61 (0.49–0.75)	6.53E–06	
			Meta-analysis[d]							0.56 (0.51–0.61)	2.31E–38	0.84
rs9277535	*HLA-DPB1*	A/G	GWAS second stage	58	254	294	230	619	418	0.59 (0.51–0.69)	1.78E–12	
			First replication	26	102	144	49	132	91	0.54 (0.42–0.69)	1.96E–06	
			Second replication	68	264	376	227	604	445	0.56 (0.48–0.64)	1.81E–16	
			Third replication	29	136	139	107	273	155	0.56 (0.46–0.69)	6.52E–08	
			Meta-analysis[d]							0.57 (0.52–0.62)	6.34E–39	0.85

Odds ratio and *P* values for independence test were calculated by the Mantel-Haenszel method.
[a]Odds ratio of minor allele from two-by-two allele frequency table. [b]*P* values of Pearson's χ^2 test for allele model. [c]Result of Breslow-Day test. [d]Meta-analysis of all four studies.

Table 2 Haplotype analysis

No.	Haplotype[a]	Frequency (cases)	Frequency (controls)	P^b	OR[b] (95% CI)
1	GG-*DPA1*0202*-TCG-*DPB1*0501*-GAGATT	0.428	0.347	5.79E–06	1.45 (1.16–1.81)
2	AA-*DPA1*0103*-CCA-*DPB1*0201*-AGTGCC	0.165	0.192	0.052	Reference
3	GG-*DPA1*0201*-TCG-*DPB1*0901*-GGGGTC	0.129	0.124	0.642	1.21 (0.91–1.61)
4	AA-*DPA1*0103*-CTA-*DPB1*0402*-AGTGCC	0.042	0.096	6.00E–08	0.52 (0.35–0.75)
5	AA-*DPA1*0103*-CCA-*DPB1*0401*-AGTGCC	0.018	0.038	0.002	0.57 (0.33–0.96)
6	GG-*DPA1*0202*-TCG-*DPB1*0301*-GGGGTC	0.036	0.018	0.002	2.31 (1.39–3.84)
7	GG-*DPA1*0202*-TCG-*DPB1*0202*-AGTGCC	0.020	0.027	0.257	0.88 (0.51–1.52)
8	GG-*DPA1*0202*-TCG-*DPB1*0201*-AGTGCC	0.022	0.024	0.662	0.97 (0.57–1.65)
9	GG-*DPA1*0201*-TCG-*DPB1*0501*-GAGATT	0.029	0.018	0.057	1.81 (1.06–3.08)
10	GG-*DPA1*0201*-TCA-*DPB1*1301*-GGTGCC	0.022	0.016	0.172	1.69 (0.95–3.03)
11	AA-*DPA1*0103*-CTG-*DPB1*0301*-GGGGTC	0.011	0.016	0.246	0.74 (0.36–1.53)
12	GG-*DPA1*0201*-TCG-*DPB1*1401*-GGGGTC	0.012	0.012	0.877	1.25 (0.61–2.53)

Controls of the first stage and cases of the second stage were analyzed.
[a]Haplotypes consisting of rs2595309, rs3077, *HLA-DPA1*, rs2301220, rs9277341, rs3135021, *HLA-DPB1*, rs9277535, rs10484569, rs3128917, rs2281388, rs3117222 and rs9380343 are shown. [b]*P* values, odds ratios and its 95% confidence intervals of each haplotype were calculated as described in the Methods.

regression model. Thus, our findings clearly indicate that hepatitis B is associated with variants in the *HLA-DP* loci.

A number of reports have described association of several *HLA* and non-*HLA* genes with persistent HBV infection[12,15], but their results were not consistent among the studies, and none of them indicated a possible involvement of the *HLA-DP* locus. This study is the first GWAS to investigate host genetic factors associated with chronic hepatitis B. One genome-wide linkage analysis using 318 microsatellite markers in the Gambian population suggested that the chromosome 21q22 region contains a susceptibility locus for persistent HBV infection[16]. However, our GWAS analysis failed to support this result, possibly owing to ancestry differences or different modes of viral transmission (the vertical transmission in Japan versus the horizontal transmission in Gambia).

To investigate the correlation between the incidence of hepatitis B infection and these polymorphisms, we evaluated the frequencies of rs3077 and rs9277535 in 11 different HapMap3 populations (**Supplementary Table 6** online). Our association analysis indicated that A alleles at both rs3077 and rs9277535 were associated with protective effects for chronic hepatitis B. Notably, the frequencies of these two alleles were lower in Asian and African populations, especially in the Chinese population, compared with European and Central American populations. Although disease prevalence is not determined solely by genetic factors, the findings presented in our manuscript suggest that genetic factors might exert substantial influence on the prevalence of infectious disease.

Antigen presentations on HLA class II molecules to CD4-positive helper T cells and on class-I molecules to CD8-positive cytotoxic T cells are considered to be critical for the immune response against exposure to HBV. Although cytotoxic T cells are suspected to have major roles in viral clearance, helper T cells are also essential in the immune response to acute infections[17]. HLA-DPs have a structure similar to other classical HLA class II molecules, but their roles in the immune response have not been well characterized, except the association with berylliosis[18]. The 11 SNPs we found showing strong association with chronic HBV infection were in very strong LD with *HLA-DP* alleles. Because the subsequent haplotype analyses identified significant association of chronic hepatitis B with haplotypes containing the *HLA-DPA1* and *HLA-DPB1* genes, we suspected that variations in HLA-DP molecules would affect the ability for antigen presentation of HLA class II molecules on immune cells and result in weak (or no) immune response and persistent HBV infection. A previous report that implicated *HLA-DPA1**0103 and *DPB1**0402 to be candidate predictive factors for antibody production after HBV vaccination[19] supports this hypothesis. It should be noted that the lack of information regarding exposure to HBV for each control might underestimate the effect size obtained in this study but does not inflate the type 1 error rate.

In summary, we have demonstrated that genetic variants in the *HLA-DP* genes are strongly associated with chronic hepatitis B in the Asian population. Considering the function of HLA-DP molecules, our findings suggest that antigen presentation on HLA-DP molecules might be critical for virus elimination and have an important role in the pathogenesis of chronic hepatitis B. An understanding of the molecular mechanism by which *HLA-DP* variants confer risk of chronic hepatitis B should shed light on its pathogenesis and facilitate development of new therapies for treatment of the disease and prevention of disease progression.

METHODS

Samples. Characteristics of each cohort group are shown in **Supplementary Table 1**. Case and control samples used in this study for the Japanese population were obtained from the BioBank Japan at the Institute of Medical Science, the University of Tokyo[20], except case samples of the second replication and control samples of the first stage of the GWAS. From the registered samples in BioBank Japan, we selected individuals that were clinically diagnosed as having chronic hepatitis B. The diagnosis of chronic hepatitis B was conducted based on HBsAg-seropositivity and elevated serum aminotransferase levels for more than six months according to the guideline for diagnosis and treatment of chronic hepatitis (see URLs section below). The control groups consisted of 2,821 individuals that were registered in BioBank Japan as subjects with diseases other than chronic hepatitis B. Subjects who were positive for HBsAg were excluded from the controls. We obtained 934 Japanese control DNAs in the first stage from volunteers in the Osaka-Midosuji Rotary Club, Osaka, Japan. Case samples for the second replication cohort ($n = 718$, RIKEN) were collected at Toranomon Hospital as well as at hospitals participating in the Hiroshima Liver Study Group (for a list of doctors participating in this study group, see URLs section below). Cases and controls for the Thai replication study ($n = 308$ and 546, respectively) were collected at Ramathibodi Hospital, Mahidol University, Thailand. The diagnosis of chronic hepatitis B was based on HBsAg-seropositivity and elevated serum aminotransferase levels. All participants provided written informed consent. This research project was approved by the ethical committees at the Institute of Medical Science, the University of Tokyo, the Center for Genomic Medicine (formerly SNP Research Center), RIKEN and Ramathibodi Hospital, Mahidol University.

SNP genotyping. We applied the two-stage approach as described previously[21]. For the first stage, we genotyped 188 individuals with chronic hepatitis B and 934 controls using the Illumina HumanHap550v3 Genotyping BeadChip. After excluding nine cases with call rate of <0.98, we applied SNP quality control (call rate of ≥ 0.99 in both cases and controls and P value of Hardy-Weinberg equilibrium test of $\geq 1.0 \times 10^{-6}$ in controls): 499,544 SNPs on autosomal chromosomes passed the quality control filters and were further analyzed. Among the SNPs analyzed in the first stage, we selected the top 12,000 SNPs showing the smallest P values for the second stage. SNPs with minor allele frequency (MAF) of ≤ 0.1 in both case and control samples were excluded from the further analysis. In the second stage, we genotyped an additional panel of 616 cases using an