TITLE PAGE

Title

Genetic risk of hepatocellular carcinoma in patients with hepatitis C virus: a case control study.

Short running title

Genetic risk of hepatocellular carcinoma

Author Name

Takeshi Tomoda[1]*, Kazuhiro Nouso[2], Akiko Sakai[3], Mamoru Ouchida[3], Sayo Kobayashi[1], Koji Miyahara[1], Hideki Onishi[2], Shinichiro Nakamura[1], Kazuhide Yamamoto[1], Kenji Shimizu[3]

Address:

[1] Department of Gastroenterology and Hepatology, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences

[2] Department of Molecular Hepatology, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences

[3] Department of Molecular Genetics, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences

*Corresponding author

Takeshi Tomoda, M.D.

Department of Gastroenterology and Hepatology,

Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences, 2-5-1 Shikata-cho, Kita-ku, Okayama City, Okayama, 700-8558, Japan

Tel:+81-86-235-7219 Fax: +81-86-225-5991

E-mail: tomotake79@yahoo.co.jp

Abstract

Backgroud and Aim: Chronic hepatitis C virus (HCV) infection is well known risk factor for hepatocellular carcinoma (HCC). The aim of this study is to elucidate the genetic risk of development and recurrence of HCC in patients with HCV.

Methods: A total of 468 patients with HCV, including 265 with HCC were enrolled. We genotyped 89 SNPs in 81 genes expected to influence hepatocarcinogenesis using the iPLEX assay. Risk of HCC was clarified by stratifying patients into risk groups based on the multiplied Odds Ratio (MOR) for SNPs associated with HCC, and the cumulative effects on the development and reccurence of HCC were analyzed.

Results: Six SNPs associated with risk of HCC were identified (OR range: 0.29~1.76). These included novel SNPs for hepatocarcinogenesis with HCV CCND2 rs1049606, RAD23B rs1805329, CEP164 rs573455, and GRP78rs430397 in addition to the known SNPs MDM2 rs2279744 and ALDH2 rs671. MOR analysis revealed that the highest risk group exerted about a 19 fold higher relative OR compared to the lowest risk group ($p = 1.08 \times 10^{-5}$). Predicted 10 year HCC risk ranged from 1.7% to 96% depending on the risk group and the extent of fibrosis. Recurrence-free survival of RFA-treated HCC in the high risk group (n=53) was lower than that of low risk group (n=58, $p = 0.038$).

Conclusion: SNPs of CCND2, RAD23B, GRP78, CEP164, MDM2, and ALDH2 genes were significantly associated with development and recurrence of HCC in Japanese patients with HCV.

Keywords: single nucleotide polymorphism; hepatocellular carcinoma; chronic hepatitis C

Introduction

More than 170 million people worldwide are estimated to have chronic hepatitis C virus (HCV) infection. The most important sequelae of chronic HCV infection are progressive liver fibrosis leading to cirrhosis and hepatocellular carcinoma (HCC), the latter responsible for significant morbidity and mortality throughout the world.[1-3] Alcohol intake, older age at time of infection, male sex, and co-infection with hepatitis B virus accelerate disease progression in HCV- infected patients[4-6], but do not fully account for the development of HCC.

As with many cancers, variants of genes involved in multistage carcinogenesis may determine an individual's susceptibility to developing HCC. Single nucleotide polymorphisms (SNPs) are the most common type of genomic sequence variation and are thought to be associated with population diversity, susceptibility to disease, and individual response to drug treatment.[7] Many SNPs are silent, with no direct effect on gene products, but by virtue of linkage disequilibrium existing across the human genome they can be used as genetic markers to locate nearby functional variants that contribute to disease. SNPs may also have functional consequences if they affect coding or regulatory (usually promoter) regions of genes. Information accumulated from numerous studies on the association between cancer risk and SNPs in selected candidate genes may shed light on the molecular and genetic basis of the polygenic nature of cancer.

We performed a search for SNPs in candidate genes associated with

susceptibility to the development of HCC. A total of 88 SNPs in 81 genes were examined in Japanese patients with chronic HCV infection. We identified two previously−reported and four novel variant SNPs as significant risk factors for incident HCC among patients with chronic HCV infection. The six SNPs were also associated with recurrence of HCC.

Methods

Study Subjects

This case-control study included 468 Japanese patients with chronic HCV infection who were admitted to Okayama University Hospital or Kagawa Prefectural Central Hospital in Japan between January 2004 and December 2009. Patients comprised 265 with HCC and 203 without HCC. Chronic HCV infection was judged by a positive test for HCV antibody. We excluded patients who tested positive for the hepatitis B surface antigen. Patients with HCC were newly diagnosed and the HCC were previously untreated. Diagnosis of HCC was made by several imaging modalities, including angiography, computed tomography, and magnetic resonance imaging, or by tumor biopsy. Diagnostic criteria for HCC via imaging were based on previous reports of hyper-attenuation at the arterial phase, hypo-attenuation at the portal phase in dynamic computed tomography (CT) or magnetic resonance imaging (MRI), and tumor staining on angiography. According to guidelines of the American Association for the Study of Liver Disease, we confirmed HCC diagnosis using at least two dynamic imaging modalities. [8] Nodules without positive imaging were histologically confirmed as HCC via ultrasound-guided, fine-needle biopsy. Any patients imbibed over 80 g/day alcohol for longer than 10 years were considered to have a positive history of alcohol abuse. Age, gender, and information on clinical status were obtained for each patient at the time of whole-blood collection. Interferon had been administered to 61 (23%) cases and 49 (24%) controls,

with 12 (5%) and 11 (5%) sustained virologic responders. Informed consent was obtained from all subjects. The study protocol conformed to the ethical guidelines of the World Medical Association Declaration of Helsinki and was approved by The Bioethics Committee of Okayama University Medical School.

Gene selection

We selected 88 SNPs in 81 candidate genes considered to play a possible role in carcinogenesis by virtue of their ability to modify cell growth, hepatic inflammation, and/or hepatocyte apoptosis. (Table 1) The candidate sites include genes related to growth factors, growth factor receptors, cytokines/chemokines, cytokine/chemokine receptors, apoptosis, tumor suppression, DNA repair, cell-cycle regulation, metabolism, cell-cell interaction, and chromosome segregation. Most of them have been reported to be associated with carcinogenesis as well as HCC susceptibility. [9-31] SNPs of the selected genes were extracted from the Japanese Single Nucleotide Polymorphisms database (http://snp.ims.u-tokyo.ac.jp), a database for SNPs found in the Japanese population, and from National Center for Biotechnology Information.

Genotyping

Genomic DNA was extracted from peripheral blood leukocytes using a QIAamp DNA blood mini kit (Qiagen, Hilden, Germany) according to the

instructions of the manufacturer. SNPs were genotyped using Sequenom® MassARRAY technology (Sequenom®, San Diego, CA, USA). The iPLEX™ assay was conducted according to the manufacturer's instructions using 20 ng of genomic DNA after classification into four groups of multiplex analyses. The primers for amplification and extension were designed using Mass ARRAY Assay Design v.3.1 software; primer sequences are available from the authors upon request. Briefly, DNA was amplified using PCR and the unincorporated nucleotide triphosphates were deactivated by phosphatase treatment shrimp alkaline phosphatase. A single base primer extension step was performed and allele–specific extension products of different masses were quantitatively analyzed using matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI–TOF MS).

Treatment and follow-up for HCC

Among 265 HCC patients, 137 received radiofrequency ablation (RFA) as their first treatment modality, of whom 111 met the following inclusion criteria: HCC with size ≤3 cm and without extrahepatic metastasis, tumor number 3 or less, and Child-Pugh grade A or B. After RFA, all patients underwent dynamic computed tomography (CT) or magnetic resonance imaging (MRI) and complete ablation was confirmed. The patients were followed regularly using abdominal ultrasound examination, CT, or MRI every 3 months. All patients were followed until death or their last hospital visit.

Statistical Analysis

Differences between cases and controls in the distributions of demographic characteristics were tested using Student's t-test or the χ^2 test. Tests for Hardy–Weinberg equilibrium were performed for each SNP separately among control subjects using Fisher's exact test. Differences in allele frequencies between patients with and without HCC were tested for each SNP using the χ^2 test. Odds ratios (OR) and 95% confidence intervals (CI) were calculated using the major allele as reference. Differences in genotype frequencies were tested using dominant, recessive, and overdominant genetic models for each SNP. OR and 95% CI were estimated via unconditional logistic regression with a multiplicative model adjusting for age, gender, drinking history, and Child-Pugh grades (SPSS Ver.12.0, SPSS Inc, Tokyo, Japan).

Permutation tests with 10,000 reiterations were performed using Haploview (http://www.broad.mit.edu/mpg/ haploview/) and declared statistically significant with $p<0.05$.

We tested the combined effects of the six SNPs associated with HCC development by calculating a multiplied odds ratio (MOR) separately for each individual patient. The MOR is the product of ORs for all risk genotypes detected among the six SNPs in a patient, with adjusted ORs based on the best-fitting genetic models from single-SNP analyses. When calculating the MOR , we multiplied the risk ORs of six SNPs. The MOR was

then categorized into five groups based on quintiles: very low, low, moderate, high, and very high risk. The performance of each cutoff was determined in terms of discriminatory ability. ORs, CIs, and p values for each MOR risk category were estimated using the very low risk group (lowest quintile of MOR) as the reference category.

It has been reported that among untreated patients the annual incidence of HCC increases with degree of liver fibrosis, being 0.45% among patients with liver fibrosis stage 0 or 1, 1.99% with stage 2, 5.34% with stage 3, and 7.88% with stage 4 (cirrhosis).[32] Based on the data and likelihood ratio (LR) of each MOR quintile risk category, we calculated the predicted 10-year absolute risk for HCC incidence (%) in each group according to stage of liver fibrosis as following.

$$\text{10-year risk} = \frac{\text{LR} \times \text{Odds}}{1 + \text{LR} \times \text{Odds}} \times 100$$

In this formula, Odds were converted from 10-year risk of HCC that was calculated from the reported annual incidence of HCC at each fibrosis stages.

In the analysis of HCC recurrence after RFA, we combined the five MOR risk categories into two groups: one including very low, low, and moderate risk (n = 58), the other combining high and very high risk (n = 53). Cumulative recurrence rates were estimated using the Kaplan-Meier method and compared using the log rank test. Because early development of HCC is thought to involve pre-existing intra-hepatic metastasis, we defined recurrence of HCC as a new legion that developed

more than three months after treatment, with the starting date of follow-up for tumor recurrence being the day when all tumors were ablated by RFA.

Results

Patient characteristics are shown in Table 2. There was no significant difference between HCC cases and controls in terms of alcohol intake, total bilirubin, prothrombin time, alanine aminotransferase (ALT) and IFN therapy. Age, proportion of males, and proportion of Child-Pugh grade B were higher in patients with HCC than in patients without HCC, whereas serum albumin level and platelet count were lower in patients with HCC.

SNPs related to HCC

Among the 88 SNPs, we selected ones with minor allele frequency ≥10%, genotyping success rate ≥95%, and no evidence of deviation from Hardy–Weinberg equilibrium ($p<0.05$) in control subjects. Seventy-five SNPs satisfied these criteria, of which ten showed a positive association with HCC incidence (significant crude OR): CCND2 rs1049606, RAD23B rs1805329, miR-146a rs2910164, GRP78 rs430397, MDM2 rs2279744, ALDH2 rs671, TP53 rs1042522, WRN rs1801195, XPC rs2228000, and CEP164 rs573455. To evaluate the effects of these polymorphisms on the occurrence of HCC, unconditional logistic regression analysis was performed with adjustment for age, gender, drinking history, and Child-Pugh grade. As further validation, permutation tests for allele frequencies demonstrated significant differences with six SNPs: CCND2 rs1049606, RAD23B rs1805329, MDM2 rs2279744, ALDH2 rs671, CEP164 rs573455, and GRP78 rs430397 (Table 3). After determining the best fitting genetic model for each variant (Table 4), CEP164 rs573455 had the

strongest association with HCC; adjusted OR for the CC genotype was 0.29 (95% CI: 0.15-0.56, $p=1.9 \times 10^{-4}$) using TT+TC as the reference.

Prediction of HCC development by MOR

MOR was calculated for the above-mentioned six SNPs to assess the cumulative effects on HCC incidence after inverting protective ORs (OR<1) (equivalently, using the OR for the reference genotype compared with risk genotype as reference). The risk genotype and risk OR for each SNP is shown in Table 4. Calculated MOR ranged from 1.0 to 71.7 and was categorized into five groups using cutpoints 6, 13, 24, and 70. The difference in the relative OR of HCC incidence among the risk categories was highly significant. Compared with the very low risk group, relative OR of HCC was high in the very high risk group (OR=19.1, 95% CI: 3.95-92.6, $p=1.08\times10^{-5}$) (Table 5). Table 5 also presents predicted cumulative incidence of HCC over 10 years according to stage of liver fibrosis in each MOR category. Ten-year predicted cumulative incidence ranged widely from 1.67 in the very low risk group of F0 or 1 to 96.2% in the very high risk group of F4. Within any particular fibrosis stage, the predicted incidence differed by anywhere from 1.7 to 15 fold among MOR categories, with the widest range in the lowest fibrosis stages. Even with fibrosis stage 0 or 1, predicted 10-year cumulative HCC incidence is 24.5% in the very high risk MOR group. Interferon (IFN) therapy was conducted in 61 cases (23%) before development of HCC. Twelve (5%) were sustained virologic responders and 49 (18%) were non-responders. No

difference in MOR was observed between sustained virological responders and non-responders.

Impact of MOR on Recurrence-Free Survival of RFA-treated HCC patients

We compared the recurrence-free survival of RFA-treated patients between high risk and low risk groups. Characteristics of the 111 patients are shown in Table 6. The high risk group was younger, but there was no significant difference in sex, tumor size, tumor number, alanine aminotransferase(ALT), total bilirubin, alubumin, platelet, prothrombin time, or alpha-fetoprotein. The median follow-up period was 1043 days (range, 176-2533 days). The 3-year cumulative recurrence rates were 77% in the high risk group and 59% in the low risk group. Patients in the high risk group had higher recurrence rates than those in the low risk group ($p = 0.038$) (Figure 1).

Discussion

We studied 89 SNPs in 81 genes that were expected to be associated with hepatocarcinogenesis. Putative genetic markers for susceptibility to hepatocarcinogenesis were identified in patients with chronic HCV infection. Six SNPs in six different genes were identified as being associated with HCC: CCND2 rs1049606, RAD23B rs1805329, MDM2 rs2279744, ALDH2 rs671, GRP78 rs430397, and CEP164 rs573455. To our knowledge, the findings of associations linking CCND2 rs1049606, RAD23B rs1805329, and CEP164 rs573455 with HCC are novel. The relationships of polymorphisms MDM2 rs2279744 and ALDH2 rs671 with HCC confirm previous results independently reported by other groups in Japan [30, 31, 33, 34].

SNP rs1805329 (Ala249Val) is located in the RAD23B gene. The RAD23B protein is crucial in recognition and initiation of global genomic repair (GGR). SNP rs573455 (Gln1119Arg) is located in the CEP164 gene, which encode a centriole appendage protein and also associated with DNA repair. [35, 36] SNP rs1049606 (T-171C) is located in 5'-UTR of the CCND2 gene. The protein encoded by this gene is involved in phosphorylation of the tumor-suppressor protein Rb; abnormal levels of CCND2 are associated with poor prognosis in gastric cancer and intrahepatic recurrence in HCC. [37, 38] SNP rs430397 (G/A) is located in intron 5 of the GRP78 gene and adjoins the 3'-end of the intron. The GRP78 pathway, one of the most important responders to disease-associated stresses [39], demonstrates high correlation between expression and cancer progression, has

anti-apoptotic function, and leads to drug resistance in HCC.[40, 41] MDM2 (rs2279744) in the promoter region of the MDM2 gene, a negative regulator of p53, is associated with accelerated tumor formation in both hereditary and sporadic cancers in humans. [42] SNP rs671 (Glu504Lys) is located in the ALDH2 gene. ALDH2 is a key enzyme in the elimination of acetaldehyde. [43] Our analysis revealed that heterozygotes for rs671 were associated with increased risk of HCC development; a similar result has been reported for esophageal cancer.[44] All six genes are presumed to be associated with carcinogenesis. However, functional effects of the four SNPs excluding ALDH2 and MDM2 have yet to be fully elucidated.

We calculated joint effects on HCC development for the six polymorphisms by multiplying their separate odds ratios for each patient. Although gene-gene interaction is ignored using MOR, even if interaction exists the effects should influence cases and controls similarly. Furthermore, risk can be evaluated in greater detail by calculating the MOR of two or more genes compared with a single gene. Previously, many studies demonstrated cumulative effects of SNPs in detecting the high risk group.[45, 46] Moreover, the high risk group may be refined by combining MOR with fibrosis stage, an indicator of carcinogenesis. Yet even in the same fibrosis stage, the variation in predicted incidence of HCC can range anywhere from 1.7 to 15 fold. If a patient has mild fibrosis (F0/F1/F2) but is in the high risk group based on MOR, predicted HCC incidence may be higher than for a patient in an advanced stage of fibrosis (F3/F4) with lower risk based on MOR. According to the evidence-based

clinical practice guidelines for HCC in Japan,[47] screening is recommended once every 6 months for the high risk group (patients with chronic hepatitis B, C or with cirrhosis) and once every 3-4 months for the very high risk group (patients with cirrhosis type B or C). By adding the concept of MOR, we can further refine the risk group and predict the incidence risk of HCC for each stage of fibrosis according to the risk categories stratified by MOR. Our data suggest that MOR may occasionally be high even in patients whose fibrosis stage is 0 or 1. Such patients are nevertheless at high risk of HCC, which indicates that SNP analysis complements other laboratory tests in identifying high-risk patients.

Moreover, we found that the tumor recurrence rates following RFA therapy differed between patients in the high risk and low risk groups based on MOR category. Those data increased the reliability of MOR in this study and suggest that we could predict not only risk of HCC development but also risk of recurrence. By analyzing the SNPs, we could pay more attention during the surveillance of these high risk patients, and might achieve an early diagnosis of HCC.

Several limitations of the present study should be noted. Histological examination, which is required for precise evaluation of liver fibrosis stage, was not performed on all of the non-cancerous tissues, although we used Child-Pugh grade as a parameter of liver function. This study was lacking the data of HCV viral load or genotype which might be the risk factors of HCV related HCC. In addition, the study design was

retrospective with a small number of patients. A prospective study with larger sample is needed to confirm our results. Moreover, validation study is needed to confirm the conclusion.

Six SNPs：CCND2 rs1049606, RAD23B rs1805329, MDM2 rs2279744, ALDH2 rs671, CEP164 rs573455, and GRP78 rs430397, are associated with risk of HCC among Japanese patients with chronic HCV infection. We could predict the absolute risk of HCC among HCV-related hepatitis patients by analyzing cumulative effects of these six SNPs using multiplied OR. The data were also effective for predicting patient's prognosis.

Conflict of interests

There is no conflict of interest..

Acknowledgments

This study makes use of data generated by the SNP Research Centre of RIKEN and the University of Tokyo who have performed the sample collection, genotyping, bioinformatics and data quality control. We appreciate Dr. J.B. Cologne for his excellent advises on biostatistical analyses and on the manuscript preparation.

Financial Support

This study was supported by grants in aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan to S.N. (22590736) and to K.S. (22300346).

References

1. Liang TJ, Rehermann B, Seeff LB, et al. **Pathogenesis, natural history, treatment, and prevention of hepatitis C**. *Ann Intern Med*. 2000; 132: 296-305..
2. Shiratori Y, Imazeki F, Moriyama M, et al. **Histologic improvement of fibrosis in patients with hepatitis C who have sustained response to interferon therapy**. *Ann Intern Med*. 2000; 132: 517-524.
3. Yoshida H, Arakawa Y, Sata M, et al. **Interferon therapy prolonged life expectancy among chronic hepatitis C patients**. *Gastroenterology*. 2002; 123:483-491.
4. Yeh SH, Chen PJ, Shau WY, et al. **Chromosomal allelic imbalance evolving from liver cirrhosis to hepatocellular carcinoma.** *Gastroenterology*. 2001; 121: 699-709.
5. Cerny A, Chisari FV. **Pathogenesis of chronic hepatitis C: immunological features of hepatic injury and viral persistence**. *Hepatology*. 1999; 30: 595-601.
6. Liang TJ, Rehermann B, Seeff LB,et al. **Pathogenesis, natural history, treatment, and prevention of hepatitis C**. *Ann Intern Med*. 2000; 132: 296-305.
7. Shastry BS. **SNP alleles in human disease and evolution**. *J Hum Genet* .2002; **47**: 561-566.
8. Bruix J, Sherman M. **Management of hepatocellular carcinoma**. *Hepatology*. 2005; 42: 1208-1236.
9. Homann N, Stickel F, Konig IR, et al. **Alcohol dehydrogenase 1C*1**

allele is a genetic marker for alcohol-associated cancer in heavy drinkers. *Int J Cancer*. 2006; 118: 1998-2002.

10. Akkiz H, Bayram S, Bekar A, et al. **Relationship between functional polymorphism in the Aurora A gene and susceptibility of hepatocellular carcinoma**. *J Viral Hepat*. 2010; 17: 668-674.
11. Akkiz H, Bayram S, Bekar A, et al. **Cyclin D1 G870A polymorphism is associated with an increased risk of hepatocellular carcinoma in the Turkish population: case-control study**. *Cancer Epidemiol.* 2010; 34: 298-302.
12. Xiong XD, Fang JH, Qiu FE,et al. **A novel functional polymorphism in the Cdc6 promoter is associated with the risk for hepatocellular carcinoma**. *Mutat Res.* 2008; 643: 70-74.
13. Wang XQ, Luk JM, Garcia-Barcelo M, et al. **Liver intestine-cadherin (CDH17) haplotype is associated with increased risk of hepatocellular carcinoma**. *Clin Cancer Res.* 2006; 12: 5248-5252.
14. Kato N, Ji G, Wang Y, et al. **Large-scale search of single nucleotide polymorphisms for hepatocellular carcinoma susceptibility genes in patients with hepatitis C**. *Hepatology*. 2005; 42: 846-853.
15. Hu L, Liu J, Chen X, et al. **CTLA-4 gene polymorphism +49 A/G contributes to genetic susceptibility to two infection-related cancers-hepatocellular carcinoma and cervical cancer**. *Hum Immunol* . 2010; 71: 888-891.
16. Dong X, Zhou G, Zhai Y, et al. **Association of DLC1 gene polymorphism with susceptibility to hepatocellular carcinoma in**

Chinese hepatitis B virus carriers. *Cancer Epidemiol* .2009; 33: 265-270.

17. Tanabe KK, Lemoine A, Finkelstein DM, et al. **Epidermal growth factor gene functional polymorphism and the risk of hepatocellular carcinoma in patients with cirrhosis**. *JAMA*. 2008; 299: 53-60.

18. Zhai Y, Zhou G, Deng G, *et al*: **Estrogen receptor alpha polymorphisms associated with susceptibility to hepatocellular carcinoma in hepatitis B virus carriers**. *Gastroenterology.* 2006;130: 2001-2009.

19. Chen YL, Tseng HS, Kuo WH, et al. **Glutathione S-Transferase P1 (GSTP1) gene polymorphism increases age-related susceptibility to hepatocellular carcinoma**. *BMC Med Genet*. 2010; 11: 46.

20. Couvert P, Carrie A, Paries J, et al. **Liver insulin-like growth factor 2 methylation in hepatitis C virus cirrhosis and further occurrence of hepatocellular carcinoma**. *World J Gastroenterol*. 2008; 14: 5419-5427.

21. Wang Y, Kato N, Hoshida Y, et al. **Interleukin-1beta gene polymorphisms associated with hepatocellular carcinoma in hepatitis C virus infection**. *Hepatology*. 2003; 37: 65-71.

22. Tseng LH, Lin MT, Shau WY, et al. **Correlation of interleukin-10 gene haplotype with hepatocellular carcinoma in Taiwan**. *Tissue Antigens*. 2006; 67: 127-133.

23. Xu Y, Liu L, Liu J, *et al*. **A potentially functional polymorphism in the promoter region of miR-34b/c is associated with an increased**

risk for primary hepatocellular carcinoma. *Int J Cancer*. 2011; 128: 412-417.

24. Xu T, Zhu Y, Wei QK, et al. **A functional polymorphism in the miR-146a gene is associated with the risk for hepatocellular carcinoma**. *Carcinogenesis*. 2008; 29: 2126-2131.

25. Mu LN, Cao W, Zhang ZF, et al. **Methylenetetrahydrofolate reductase (MTHFR) C677T and A1298C polymorphisms and the risk of primary hepatocellular carcinoma (HCC) in a Chinese population**. *Cancer Causes Control*. 2007; 18: 665-675.

26. Kwak SY, Kim UK, Cho HJ, et al. **Methylenetetrahydrofolate reductase (MTHFR) and methionine synthase reductase (MTRR) gene polymorphisms as risk factors for hepatocellular carcinoma in a Korean population**. *Anticancer Res*. 2008; 28: 2807-2811.

27. Segat L, Milanese M, Pirulli D, et al. **Secreted protein acidic and rich in cysteine (SPARC) gene polymorphism association with hepatocellular carcinoma in Italian patients**. *J Gastroenterol Hepatol*. 2009; 24: 1840-1846.

28. Kim YJ, Lee HS, Im JP, et al. **Association of transforming growth factor-beta1 gene polymorphisms with a hepatocellular carcinoma risk in patients with chronic hepatitis B virus infection**. *Exp Mol Med*. 2003; 35: 196-202.

29. Borentain P, Gerolami V, Ananian P,et al. **DNA-repair and carcinogen-metabolising enzymes genetic polymorphisms as an independent risk factor for hepatocellular carcinoma in Caucasian**

liver-transplanted patients. *Eur J Cancer.* 2007; 43: 2479-2486.

30. Dharel N, Kato N, Muroyama R,et al. **MDM2 promoter SNP309 is associated with the risk of hepatocellular carcinoma in patients with chronic hepatitis C**. *Clin Cancer Res.* 2006; 12: 4867-4871.

31. Munaka M, Kohshi K, Kawamoto T, et al. **Genetic polymorphisms of tobacco- and alcohol-related metabolizing enzymes and the risk of hepatocellular carcinoma**. *J Cancer Res Clin Oncol.* 2003; 129: 355-360.

32. Yoshida H, Shiratori Y, Moriyama M,et al. **Interferon therapy reduces the risk for hepatocellular carcinoma: national surveillance program of cirrhotic and noncirrhotic patients with chronic hepatitis C in Japan. IHIT Study Group. Inhibition of Hepatocarcinogenesis by Interferon Therapy**. *Ann Intern Med.* 1999; 131: 174-181.

33. Kato S, Tajiri T, Matsukura N,et al. **Genetic polymorphisms of aldehyde dehydrogenase 2, cytochrome p450 2E1 for liver cancer risk in HCV antibody-positive japanese patients and the variations of CYP2E1 mRNA expression levels in the liver due to its polymorphism**. *Scand J Gastroenterol.* 2003; 38: 886-893.

34. Sakamoto T, Hara M, Higaki Y,et al. **Influence of alcohol consumption and gene polymorphisms of ADH2 and ALDH2 on hepatocellular carcinoma in a Japanese population**. *Int J Cancer.* 2006; 118: 1501-1507.

35. Nishi R, Okuda Y, Watanabe E,et al. **Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma**

pigmentosum group C protein. *Mol Cell Biol.* 2005; **25**: 5664-5674.

36. Graser S, Stierhof YD, Lavoie SB,et al. **Cep164, a novel centriole appendage protein required for primary cilium formation**. *J Cell Biol.* 2007; 179: 321-330.

37. Takano Y, Kato Y, Masuda M, et al. **Cyclin D2, but not cyclin D1, overexpression closely correlates with gastric cancer progression and prognosis**. *J Pathol.* 1999; 189: 194-200.

38. 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: 923-929.

39. Ji C, Shinohara M, Kuhlenkamp J, et al. **Mechanisms of protection by the betaine-homocysteine methyltransferase/betaine system in HepG2 cells and primary mouse hepatocytes**. *Hepatology.* 2007; 46 : 1586-1596.

40. Chiou JF, Tai CJ, Huang MT,et al. **Glucose-regulated protein 78 is a novel contributor to acquisition of resistance to sorafenib in hepatocellular carcinoma**. *Ann Surg Oncol.* 2010; **17**: 603-612.

41. Wang Q, He Z, Zhang J,et al. **Overexpression of endoplasmic reticulum molecular chaperone GRP94 and GRP78 in human lung cancer tissues and its significance**. *Cancer Detect Prev.* 2005; 29: 544-551.

42. Bond GL, Hu W, Bond EE,et al. **A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway**

and accelerates tumor formation in humans. *Cell*. 2004; 119: 591-602.

43. Harada S, Agarwal DP, Goedde HW,et al. **Possible protective role against alcoholism for aldehyde dehydrogenase isozyme deficiency in Japan**. *Lancet*. 1982; 2 : 827.

44. Cui R, Kamatani Y, Takahashi A, *et al*: **Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk**. *Gastroenterology*. 2009;137: 1768-1775.

45. Ripatti S, Tikkanen E, Orho-Melander M,et al. **A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses**. *Lancet* . 2010; 376: 1393-1400.

46. Zheng SL, Sun J, Wiklund F, *et al*. **Cumulative association of five genetic variants with prostate cancer**. *N Engl J Med.* 2008; 358: 910-919.

47. Makuuchi M, Kokudo N: **Clinical practice guidelines for hepatocellular carcinoma: the first evidence based guidelines from Japan**. *World J Gastroenterol* 2006; **12**: 828-829.

FIGURE LEGEND

Fig. 1 – Recurrence free survival of high risk group and low risk group. Solid line, high risk group (n = 53); dotted line, low risk group (n = 58). High risk group patients had higher recurrence rate than those in low risk group ($p = 0.038$).

Table 1: SNPs analyzed

Gene symbol	NCBI SNP ID	Gene symbol	NCBI SNP ID
ADH1B	rs1229984	IRF3	rs7251
ADH1C	rs698	ITGAV	rs2290083
ALDH2	rs671†	JUP	rs1126821
ATR	rs2227928	LPL	rs328
AURKA	rs2273535	MAD1L1	rs1801368
AXIN2	rs2240308	MDM2	rs2279744†
BARD1	rs2070094	miR-146a	rs2910164†
CASP9	rs1052571	miR-34b/c	rs4938723
CCND1	rs9344	MMP27	rs1276286
CCND2	rs1049606†	MMP9	rs17577
CDC6	rs4134994	MTHFR	rs1801133
CDH17	rs2514813 rs3214050	MTRR	rs10380
CEP110	rs10818504	NIN	rs2236316
CEP152	rs2289181	NOB1	rs3811348
CEP164	rs573455†	NSL1	rs15702
CEP192	rs578208	PCNT	rs2070425
CEP250	rs3748433	PKCI	rs481781
CEP55	rs3740370 rs2293277	POL I	rs8305
CEP57	rs644799	PTPN13	rs2230600 rs989902
CEP68	rs12611491	PTPRJ	rs1566734 rs1503185
CEP72	rs868649	RAD18	rs373572
CRHR2	rs2267716	RAD23B	rs1805329†
CTLA4	rs231775	RAG1	rs3740955
CYP1B1	rs1056836	RAPGEF6	rs1291602
CYP2C19	rs4986893	RASSF1	rs2073498
DLC1	rs621554	RASSF6	rs12507775
DUSP6	rs2279574	SCYB14	rs2237062
EGF	rs4444903	SNAI 1	rs4647958
EGFR	rs2293347 rs763317	SNM1B	rs11552449
ERBIN	rs36303	SPARC	rs2304052
ESR1	rs2077647	SRD5A2	rs523349
ETL1	rs7439869	TDG	rs4135113
EXO1	rs4149963 rs1047840	TGFB1	rs1800469
FSHR	rs6165	TP53	rs1042522†
GFRA1	rs12762746	TRAP1	rs2074805
GRP78	rs430397†	WISP3	rs1230345
GSTP1	rs1695	WRN	rs1801195†
HER2	rs1136201	XPC	rs2228000†
IGF2	rs11541372	XPG	rs17655
IL10	rs1800872	XRCC	rs25487
IL1B	rs1143627 rs16944		

† SNPs with significantly different frequencies between cases and controls.

Table 2: Characteristics of the patients

	Cases	Controls	p value *
Patients number	265	203	
Age (years)	68.4 (40-87)	57.7 (21-86)	< 0.001
Sex (male)	182 (68.7%)	100 (49.3%)	< 0.001
Alcohol >80 g/day	22 (8.3%)	13 (6.4%)	0.440
TB (mg/dL)	1.1 (0.3-27.9)	1.3 (0.3-13)	0.423
Albumin (g/dL)	3.6 (2.1-4.7)	3.9 (2.1-5.1)	<0.001
PT (%)	96 (54-146)	99 (23-152)	0.072
ALT (IU/L)	54 (10-269)	61 (9-368)	0.087
Platelet ($\times 10^4/\mu L$)	13 (2.1-41.9)	16 (1.5-34.7)	0.016
Child-Pugh grade			
A	217 (81.9%)	173 (85.2%)	0.003
B	43 (16.2%)	17 (8.4%)	
C	5 (1.9%)	13 (6.4%)	
IFN therapy	61(23%)	49(24%)	0.777

Values are median (range) or number (%).

Abbreviations: TB, total bilirubin; PT, prothrombin time; ALT, alanine aminotransferase; IFN, interferon.

* *p* values were derived from the Pearson $\chi 2$ test or student's t test.

Table 3: Alleles and genotype frequencies of six SNPs demonstrating a significant difference between cases and controls

Gene symbol Chromosomal Region	SNP ID	Alleles	MAF		OR	95% CI	Permutation *p* value†	Genotype	Case	Control	p value*
			Case (%)	Control (%)					(n)	(n)	
CCND2	rs1049606	T / C	37.5	46.6	0.69	0.53-0.90	0.0051	TT	99	64	
2p13.321								TC	133	89	0.869
								CC	33	50	0.001
RAD23B	rs1805329	C / T	17.2	25.6	0.60	0.44-0.83	0.003	CC	180	117	
9q31.2								CT	79	68	0.167
								TT	6	18	6.8×10^{-4}
GRP78	rs430397	G / A	15.1	10.0	1.58	1.06-2.36	0.034	GG	192	164	
9q33.3								GA	66	37	0.068
								AA	7	2	0.155
MDM2	rs2279744	G / T	40.9	47.7	0.76	0.58-0.99	0.038	GG	88	56	
12q15								GT	129	96	0.471
								TT	41	47	0.031
ALDH2	rs671	G / A	29	21.6	1.48	1.09-2.01	0.012	GG	132	126	
12q24.12								GA	111	60	0.004
								AA	21	13	0.244
CEP164	rs573455	T / C	38.7	45.4	0.76	0.58-0.99	0.035	TT	88	61	
11q23.3								TC	138	93	0.895
								CC	30	43	0.012

Abbreviation: MAF, minor allele frequency

* χ^2 test for 2×2 contingency table

† Permutation tests with 10,000 reiterations were performed by the Haploview program

Table 4: Best-fitting model fit to data on six SNPs with significant case-control frequency differences

Gene symbol	SNP ID	Best-Fitting Model†			Adj OR (95%CI)	p value*	Risk genotype	Risk OR
		Model	Genotype					
			Reference	Associated				
CCND2	rs1049606	Recessive	TC+TT	CC	0.50 (0.28-0.88)	0.016	TC+TT	2.00
RAD23B	rs1805329	Dominant	CC	CT+TT	0.56 (0.36-0.89)	0.013	CC	1.79
GRP78	rs430397	Dominant	GG	GA+AA	1.76 (1.04-2.96)	0.035	GA+AA	1.76
MDM2	rs2279744	Recessive	GT+GG	TT	0.50 (0.28-0.87)	0.014	GT+GG	2.00
ALDH2	rs671	Overdominant	GG+AA	GA	1.64 (1.03-2.60)	0.037	GA	1.64
CEP164	rs573455	Recessive	TC+TT	CC	0.29 (0.15-0.55)	1.9×10^{-4}	TC+TT	3.45

Abbreviation: Adj OR, adjusted odds ratio

† The best-fitting model for each SNP was determined after testing association in a series of genetic models, including dominant, recessive and overdominant models.

* Ors of genotypes and minor alleles were calculated by logistic regression adjusted for age, sex, alcohol drinking, and Child-Pugh grade and 95%CIs and p values were derived from the Wald test.

Table 5: Predicted HCC incidence calculated according to the joint effects of six SNPs

Category	MOR Range	Case	Control	OR (95% CI)	p value*	LR	10-year predicted HCC incidence (%)			
		n = 265	n = 203				F0 or 1	F2	F3	F4
		n (%)	n (%)				mean = 4.5	19.9	53.4	78.8
Very low risk	1 - 6	16 (6.0)	34 (16.7)	1.00 (Ref.)	1	0.36	1.67	8.20	29.2	57.2
Low risk	6 - 13	60 (22.6)	62 (30.5)	2.06 (1.03-4.11)	0.043	0.74	3.37	15.5	45.9	73.3
Moderate risk	13 - 24	62 (23.4)	48 (23.6)	2.74 (1.36-5.55)	0.006	0.99	4.45	19.7	53.2	78.6
High risk	24 - 70	109 (41.1)	57 (28.1)	4.06 (2.07-7.98)	3.46×10^{-5}	1.46	6.44	26.6	62.6	84.4
Very high risk	70 <	18 (6.8)	2 (1.0)	19.1 (3.95-92.6)	1.08×10^{-5}	6.89	24.5	63.1	88.8	96.2

Abbreviations: MOR, multiplied odds ratio; Ref、reference; LR, likelihood ratio; F, stage of liver fibrosis

* p values were calculated using Fisher's exact test (two-sided)

Table 6: Characteristics of RFA-treated patients

	High risk group (n = 53)	Low risk group (n = 58)	p value*
Age (years)	67.3 (40-83)	70.9 (52-85)	0.027
Sex （male）	28 (53%)	33 (57%)	0.670
Tumor size (mm)	16 (9-29)	16 (6-30)	0.606
Tumor number (single)	34 (64%)	46 (79%)	0.156
ALT (IU/L)	55 (17-235)	57 (21-164)	0.696
TB (mg/dL)	0.9 (0.4-2.2)	0.9 (0.4-1.9)	0.540
Albumin (g/dL)	3.6 (2.7-4.6)	3.7 (2.8-4.6)	0.215
Platelet count ($\times 10^4/\mu L$)	11.2 (3.1-21.5)	11.1 (4.4-30.7)	0.911
PT (%)	97.8 (71-146)	97.8 (56-140)	0.990
AFP (μg/L)	71.3 (1.9-925)	97.9 (3.2-2818)	0.635

Values are median (range) or number (%).

Abbreviations: ALT, alanine aminotransferase; TB, total bilirubin; PT, prothrombin time; AFP, α-fetoprotein

* p values were derived from either the Pearson $\chi 2$ test or student's t test.