J Gastroenterol (2016) 51:370–379
DOI 10.1007/s00535-015-1116-6

The Japanese Society of Gastroenterology

ORIGINAL ARTICLE—LIVER, PANCREAS, AND BILIARY TRACT

The impact of *PNPLA3* and *JAZF1* on hepatocellular carcinoma in non-viral hepatitis patients with type 2 diabetes mellitus

Misuzu Ueyama[1,2] · Nao Nishida[1] · Masaaki Korenaga[1] · Keiko Korenaga[1] · Erina Kumagai[1,2] · Hidekatsu Yanai[3] · Hiroki Adachi[3] · Hisayuki Katsuyama[3] · Sumie Moriyama[3] · Hidetaka Hamasaki[3] · Akahito Sako[3] · Masaya Sugiyama[1] · Yoshihiko Aoki[1] · Masatoshi Imamura[1] · Kazumoto Murata[1] · Naohiko Masaki[1] · Takumi Kawaguchi[4] · Takuji Torimura[4] · Hideyuki Hyogo[5] · Hiroshi Aikata[5] · Kiyoaki Ito[6] · Yoshio Sumida[7,8] · Akio Kanazawa[9] · Hirotaka Watada[9] · Koji Okamoto[10] · Kenjiro Honda[10] · Kazuyoshi Kon[2] · Tatsuya Kanto[1] · Masashi Mizokami[1] · Sumio Watanabe[2]

Received: 10 July 2015 / Accepted: 14 August 2015 / Published online: 3 September 2015

Abstract

Background Type 2 diabetes mellitus (T2DM) is an established independent risk factor for hepatocellular carcinoma (HCC). T2DM is associated with non-alcoholic steatohepatitis (NASH), which is a major cause of non-HBV and non-HCV-related HCC; nevertheless, it has been difficult to identify those patients with T2DM who have a high risk of developing HCC. The aim of this study was to identify genetic determinants that predispose T2DM patients to HCC by genotyping T2DM susceptibility loci and *PNPLA3*.

Methods We recruited 389 patients with T2DM who satisfied the following three criteria: negative for HBs-Ag and anti-HCV Ab, alcohol intake <60 g/day, and history of T2DM >10 years. These patients were divided into two groups: T2DM patients with HCC (DM–HCC, $n = 59$) or those without HCC (DM–non-HCC, $n = 330$). We genotyped 51 single-nucleotide polymorphisms (SNPs) previously reported as T2DM or NASH susceptibility loci (*PNPLA*3) compared between the DM-HCC and DM-non-HCC groups with regard to allele frequencies at each SNP.

Results The SNP rs738409 located in *PNPLA3* was the greatest risk factor associated with HCC. The frequency of the *PNPLA3* G allele was significantly higher among DM–HCC individuals than DM–non-HCC individuals (OR 2.53, $p = 1.05 \times 10^{-5}$). Among individuals homozygous for the *PNPLA*3 G allele ($n = 115$), the frequency of the *JAZF1* rs864745 G allele was significantly higher among DM–HCC individuals than DM–non-HCC individuals (OR 3.44, $p = 0.0002$).

M. Ueyama and N. Nishida contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00535-015-1116-6) contains supplementary material, which is available to authorized users.

✉ Masaaki Korenaga
dmkorenaga@hospk.ncgm.go.jp

[1] Present Address: The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, 1-7-1 Kohnodai, Ichikawa, Chiba 272-8516, Japan

[2] Department of Gastroenterology, Juntendo University School of Medicine, Bunkyo-Ku, Tokyo, Japan

[3] Department of Internal Medicine, National Center for Global Health and Medicine Kohnodai Hospital, Ichikawa, Chiba, Japan

[4] Division of Gastroenterology, Department of Medicine, Kurume University School of Medicine, Kurume, Japan

[5] Department of Gastroenterology and Metabolism, Hiroshima University, Hiroshima, Japan

[6] Division of Gastroenterology, Department of Internal Medicine, Aichi Medical University School of Medicine, Nagakute, Japan

[7] Center for Digestive and Liver Diseases, Nara City Hospital, Nara, Japan

[8] Department of Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, Kamigyo-ku, Kyoto, Japan

[9] Department of Metabolism and Endocrinology, Juntendo University Graduate School of Medicine, Bunkyo-Ku, Tokyo, Japan

[10] Department of Nephrology and Endocrinology, Department of Hemodialysis and Apheresis, University Hospital, The University of Tokyo, Tokyo, Japan

Conclusions *PNPLA3* and *JAZF1* were associated with non-HBV and non-HCV-related HCC development among Japanese patients with T2DM.

Keywords Type 2 diabetes mellitus · Hepatocellular carcinoma · *JAZF1* · *PNPLA3*

Abbreviations

T2DM	Type 2 diabetes mellitus
NASH	Non-alcoholic steatohepatitis
HCC	Hepatocellular carcinoma
SNPs	Single-nucleotide polymorphisms
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HBs-Ag	Hepatitis B surface antigen
anti-HCV Ab	Anti-hepatitis C virus antibody
PNPLA3	Patatin-like phospholipase domain-containing protein 3
JAZF1	Juxtaposed with another zinc finger protein 1
anti-HBc Ab	Anti-hepatitis B core antibody

Introduction

Type 2 diabetes mellitus (T2DM) is a lifestyle-related disease that affects approximately 382 million patients worldwide [1]. Recently, an association between T2DM and cancer risk, particularly hepatocellular carcinoma (HCC), has been reported. Several cohort studies in Japan revealed that the risk of liver cancer incidence was extremely high among individuals with a history of diabetes mellitus [hazard ratio (HR) 1.94, 95 % CI 1.65–2.36), even higher than that of pancreatic cancer (HR 1.85, 95 % CI 1.46–2.34) [2]. Among 820,900 participants from 97 prospective studies, diabetes mellitus was associated with death from cancer (HR 2.32, 95 % CI 2.11–2.56) with the highest HR (2.16, 95 % CI 1.62–2.88) seen for liver cancer [3]. Several clinical variables such as age, sex, and body mass index (BMI) are associated with HCC prevalence [4], and T2DM patients often exhibit risk phenotypes associated with these factors. However, T2DM patients with these clinical risk phenotypes do not often develop HCC; this observation indicates that genetic factors might influence HCC development in patients with T2DM.

Multiple reports indicate that susceptible genes associated with these comorbid diseases, such as T2DM are also associated with various cancers. Notably, single-nucleotide polymorphisms (SNPs) that are reportedly associated with susceptibility to T2DM are also associated with prostate cancer [5] and colorectal cancer [6]. Therefore, we hypothesized that T2DM and HCC may be influenced by the same susceptible genes. Many SNPs that are associated with T2DM have been identified and confirmed in multiple genome-wide association studies (GWASs) [7–18]. However, relationships between HCC development and these SNPs have not been examined.

In addition, GWASs have successfully identified the rs738409 SNP in patatin-like phospholipase domain-containing protein 3 (*PNPLA3*) (I148 M) as being associated with nonalcoholic fatty liver disease (NAFLD)/nonalcoholic steatohepatitis (NASH) [19]. Furthermore, *PNPLA3* variants are associated with not only hepatic fibrosis, but also NAFLD-related HCC in European Caucasians [20]. Because the prevalence of NAFLD/NASH is high in patients with T2DM [21], we reasoned that *PNPLA3* genotype could influence HCC development in patients with T2DM. The aim of study was to identify genetic variants that predispose T2DM patients to HCC by genotyping the SNP rs738409 in *PNPLA3* and 50 other SNPs reportedly associated with T2DM.

Methods

Patients

We recruited 389 T2DM patients, 59 with HCC and 330 without HCC from Kurume University Hospital, Hiroshima University Hospital, Aichi Medical University Hospital, Nara City Hospital, and Kohnodai Hospital in Japan. Each participant satisfied the following three inclusion criteria: (1) negative for HBs-Ag and anti-HCV Ab, (2) alcohol intake <60 g/day, (3) disease duration of T2DM >10 years. The presence of T2DM was determined on the basis of fasting blood glucose levels >126 mg/dl or hemoglobin A1c (HbA1c) >6.5 % in accordance with the diagnostic criteria for diabetes mellitus, or by the use of anti-diabetic agents. To minimize the effects of stratification on genetic associations, only individuals of Japanese ancestry were included in this study.

The 59 T2DM patients diagnosed with primary HCC were classified into the DM–HCC group. HCC was diagnosed by a combination of imaging procedures and tests for serum tumor markers such as alpha-fetoprotein and des-gamma-carboxy prothrombin. For 49 of the 59 cases, a surgical specimen or tumor biopsy was available, and in each case (49/49), the HCC diagnosis was confirmed histologically.

The 330 T2DM patients without HCC were classified into the DM–non-HCC group. DM–non-HCC patients were subsequently stratified based on liver stiffness measurements (LSMs) determined via transient elastography (Fibroscan, Echosens, Paris, France); this stratification was

potentially significant because the DM–non-HCC group may include patients with liver cirrhosis who are at a higher risk of developing HCC in the future. Among the 330 DM–non-HCC patients, 198 showed no indication of advanced hepatic fibrosis and were classified as the DM-control group. The inclusion criteria for DM-control were as follows: (1) >65 years old and (2) LSM <7 kilopascals [22] (Fig. 1).

Written informed consent was obtained from each participant. This study was conducted in accordance with provisions of the 1975 Declaration of Helsinki and approved by the Institutional Ethics Committee of National Center for Global Health and Medicine (NCGM-A-000206) and each hospital participating in the study.

DNA preparation

Genomic DNA was extracted from peripheral blood lymphocytes via the phenol chloroform DNA extraction method; purified DNA was resuspended in TE buffer. Each sample was stored at −20 °C until use.

Genotyping

The SNP rs738409 has a non-synonymous variant and is located in the third exon of *PNPLA3*; it was genotyped in each sample via TaqMan SNP genotyping assays (Applied Biosystems, Foster City, CA, USA) that were conducted with the LightCycler 480 Real-Time PCR System (Roche, Mannheim, Germany).

In addition, 58 other SNPs were selected as candidate SNPs for this study; each SNP has a documented association with T2DM in GWASs during 2003–2010, and they are located in or near the following loci: *EIF2AK4* [7], *KRT4* [7], *FAM60A* [7], *ANGPT4* [7], *SPDEF* [7], *A2BP1* [7], *Intergenic* [7], *GCKR* [7], *SLC30A8* [7–10], *HHEX* [7–10], *CDKN2B* [7–10], *IGF2BP2* [7–10], *CDKAL1* [7–10], *TCF7L2* [7–9], *KCNJ11* [7–10], *PPARG* [7–9], *WFS1* [8], *JAZF1* [8, 9], *TSPAN8/LGR5* [8, 9], *ADAMTS9* [8, 9], *CDC123/CAMK1D* [8, 9], *NOTCH2* [8, 9], *THADA* [8, 9], *FTO* [8, 9], *CALPN10* [9], *HNF1A* [9], *VEGFA* [9], *BCL11A* [9], *HNF1B* [7, 9], *KCNQ1* [7, 11, 12], *SRR* [12], *PTPRD* [12], *MAF/WWOX* [12], *ACACB* [13], *CACNA1D* [13], *CLIC5* [13], *KCNJ15* [14], *GCK1* [15], *SLC12A3* [16], *NCALD* [17], and *ELMO1* [18]. Each of the 58 SNPs was genotyped via multiplex SNP typing assay (DigiTag 2 assay [23]). Of the 58 SNPs, eight were not subject to further analysis because of a low call rate <0.95. None of the remaining 51 SNPs (including *PNPLA3* rs738409) showed Hardy–Weinberg equilibrium $p < 0.001$.

After the quality control filtering (i.e., SNP call rate ≥95 %, and HWE p value ≥0.001), the allele frequencies of the 51 SNPs (including *PNPLA3* rs738409) were compared in this analysis.

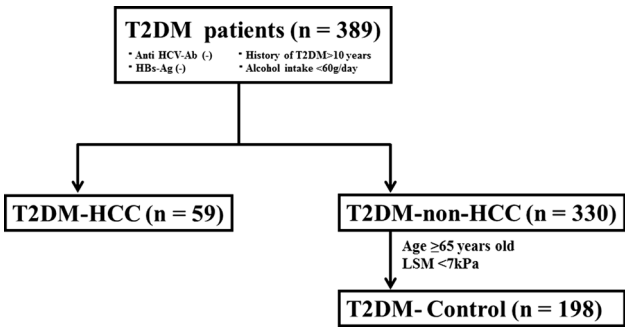

Fig. 1 Scheme of classification for the study group. The allele frequencies for each of the 51 SNPs that passed the filtering criteria were compared between the DM–HCC and DM–non-HCC groups and separately between the DM–HCC and DM-control groups

Statistical analysis

Baseline data on 19 continuous variables for each of three groups (DM–HCC, DM–non-HCC, and DM-control) are shown in Table 1 with the mean value and standard deviation; p values were calculated using the Mann–Whitney U test. Categorical values are also shown in Table 1 with the observed number of samples and the percentage of observations (%); p values were calculated with χ^2 test.

The statistical significance of associations between each of 51 SNPs and HCC was assessed via χ^2 test. To eliminate false-positive results due to multiple testing, statistical significance level was set as 0.001 (0.05/51). Subsequently, a multivariable logistic regression analysis was conducted; this analysis incorporated genotypes and biologically relevant covariates such as age, sex, and BMI that are each known to be associated with risk of progression to HCC [4].

Results

Clinical characteristics

DM–HCC patients were significantly older than DM–non-HCC patients (76.2 ± 7.5 years for DM–HCC and 70.0 ± 10.9 years for DM–non-HCC, respectively, $p < 0.0001$), and there were significantly more males in the DM–HCC group than in the DM–non-HCC group ($p = 0.0336$). The clinical characteristics related to liver injury, such as white blood cells count (WBC), platelet count (PLT), serum albumin (Alb), aspartate aminotransferase (AST), alkaline phosphatase (ALP), gamma-glutamyltranspeptidase (GGT), total bilirubin (T-Bil), FIB-4 index, LSM, ferritin, and hyaluronic acid were each

Table 1 Clinical characteristics

Characteristics	DM–HCC	DM–non-HCC	DM-control	*p* value*	*p* value**
Patients (numbers)	59	330	198		
Age (years)	76.2 ± 7.5	70 ± 10.9	74.2 ± 6.9	<0.0001	0.0157
Sex (male/female)	38/21	163/167	107/91	0.0336	0.1587
Body mass index (kg/m^2)	24.3 ± 3	24.9 ± 4.5	23.7 ± 3.7	0.3009	0.2603
WBC ($\times 10^2$/μl)	51.4 ± 20.3	61.4 ± 17.8	60.9 ± 17.3	0.0002	0.0003
Hb (g/dl)	12.7 ± 1.9	13.2 ± 1.9	13.2 ± 1.9	0.0361	0.1075
PLT ($\times 10^4$/μl)	14.1 ± 6.6	21.4 ± 6.7	22.3 ± 6.1	<0.0001	<0.0001
Albumin (g/dl)	3.8 ± 0.6	4.2 ± 0.5	4.2 ± 0.5	<0.0001	<0.0001
AST (IU/l)	50.3 ± 51.5	35.4 ± 26.2	26.2 ± 13.8	<0.0001	<0.0001
ALT (IU/l)	41.8 ± 42.7	39.9 ± 38.8	27.2 ± 18.8	0.1282	<0.0001
ALP (IU/l)	385.5 ± 167.1	264.3 ± 128	245.4 ± 95.1	<0.0001	<0.0001
GGT (IU/l)	140.1 ± 176.2	64.7 ± 99.5	41.5 ± 48.8	<0.0001	<0.0001
T-Bil (mg/dl)	1 ± 0.5	0.7 ± 0.5	0.6 ± 0.3	<0.0001	<0.0001
T-Cho (mg/dl)	174.8 ± 39.3	187.6 ± 41.4	187.3 ± 37.1	0.0749	0.0599
FBS (ng/ml)	134.2 ± 56	144.8 ± 53.5	145.9 ± 53	0.0324	0.0382
HbA1c (%)	6.8 ± 1.4	7.7 ± 1.7	7.7 ± 1.7	<0.0001	<0.0001
Cre (mg/dl)	0.8 ± 0.3	0.8 ± 0.3	0.9 ± 0.4	0.8112	0.7013
LSM (kPa)	21.4 ± 17.1	7.1 ± 6.7	4.2 ± 1.1	<0.0001	<0.0001
FIB4 index	5.2 ± 3.7	2.2 ± 1.7	1.9 ± 0.8	<0.0001	<0.0001
Hyaluronic acid (ng/ml)	294.4 ± 413	108.7 ± 173.4	69.9 ± 54.4	<0.0001	<0.0001
Ferritin (ng/ml)	249.1 ± 224.9	186.4 ± 295.8	154.2 ± 279.5	0.0077	0.0017
Anti-HBc Ab (positive/negative)	19/40	80/250	54/144	0.196	0.461
Fatty liver (presence/absence)[a]	18/41	205/125	98/100	<0.0001	0.0101
Cirrhosis (presence/absence)[b]	28/37	22/308	0/198	<0.0001	<0.0001
Liver biopsy (done/undone)	49/10	47/283	0/198		
Sulfonylurea	12 (20.3)	87 (26.4)	63 (31.8)	0.3278	0.0887
Metformin	7 (11.9)	97 (29.4)	54 (27.3)	0.0051	0.0146
Exogenous insulin	11 (18.6)	83 (25.2)	53 (26.8)	0.2822	0.2053

p values associated with continuous variables were calculated via a Mann–Whitney *U* test. *p* values associated with categorical values were calculated with χ^2 test

WBC white blood cells count, *Hb* hemoglobin, *PLT* platelet count, *Alb* serum albumin, *AST* aspartate aminotransferase, *ALT* alanine aminotransferase, *ALP* alkaline phosphatase, *GGT* gamma-glutamyltranspeptidase, *T-Bil* total bilirubin, *T-Cho* total cholesterol, *FBS* fasting blood sugar, *HbA1c* hemoglobin A1c, *Cre* creatinine, *LSM* liver stiffness measurement, *anti-HBc Ab* anti-hepatitis B core antibody

* *p* value for comparison between DM-HCC and DM-non-HCC

** *p* value for comparison between DM-HCC and DM-Control

[a] Fatty liver was diagnosed by abdominal ultrasonography or histology

[b] Cirrhosis was histologically and/or clinically diagnosed by a combination of liver biopsy, imaging analyses, and laboratory tests

significantly different between the DM–HCC and DM–non-HCC groups. There were no differences in anti-HBc Ab between the groups. No significant difference was observed in the use of sulfonylurea and exogenous insulin (Sulfonylurea: $p = 0.33$, Insulin: $p = 0.28$). The proportion of use of metformin in DM–HCC patients was significantly lower than that of DM–non-HCC group (OR 0.32, 95 % CI 0.14–0.74, $p = 0.0051$). Comparing DM–HCC to DM-control instead of DM–non-HCC, a similar tendency was observed (Table 1).

The association between HCC development and the SNPs examined in this study

Each of 58 SNPs with documented associations with T2DM were genotyped for each of the 389 Japanese T2DM patient samples. An additional SNP *PNPLA3* rs738409, which is associated with severity of nonalcoholic fatty liver disease in Japanese patients, was also genotyped in the same set of 389 samples. Quality control filtering (i.e., SNP call rate ≥95 %, and HWE *p* value ≥0.001) identified 51

Table 2 The SNPs associated with HCC in T2DM patients

SNP no.	Gene	DM–HCC ($n = 59$)	DM–non-HCC ($n = 330$)	DM-control ($n = 198$)	DM–HCC vs. DM–non-HCC		DM–HCC vs. DM-control	
					Unadjusted OR [95 % CI]	*p* value	Unadjusted OR [95 % CI]	*p* value
Allelic model								
rs738409	*PNPLA3* (G/C)	0.70/0.30	0.48/0.52	0.42/0.58	2.53 [1.66–3.87]	1.05E−05	3.22 [2.07-5.01]	1.01E−07
rs3785233	*A2BP1* (C/A)	0.24/0.76	0.16/0.84	0.14/0.86	1.63 [1.01–2.61]	0.0422	1.85 [1.11-3.08]	0.0166
rs7754840	*CDKAL1* (G/C)	0.63/0.37	0.52/0.48	0.52/0.48	1.54 [1.03–2.31]	0.0336	1.55 [1.02-2.37]	0.0407
Recessive model								
rs738409	*PNPLA3* (GG/nonGG)	0.51/0.49	0.26/0.74	0.22/0.78	2.98 [1.69–5.26]	0.0001	3.73 [2.02–6.88]	1.33E−05
rs3785233	*A2BP1* (CC/nonCC)	0.03/0.97	0.03/0.97	0.03/0.97	1.02 [0.22–4.71]	0.9823	1.12 [0.22–5.72]	0.889
rs7754840	*CDKAL1* (GG/nonGG)	0.39/0.61	0.30/0.70	0.30/0.70	1.51 [0.85–2.69]	0.1559	1.47 [0.80–2.69]	0.2108

SNPs (including *PNPLA3* rs738409) that fulfilled the filtering criteria; the allele frequencies of these 51 SNPs were used for comparisons between the DM–HCC and DM–non-HCC groups and separately between the DM–HCC and DM-control groups (Supplementary Table 1).

The *PNPLA3* rs738409 showed the strongest association of all 51 SNPs with the presence of HCC in patients with T2DM (Table 2). The allele frequency of the *PNPLA3* rs738409 G allele was significantly higher among DM–HCC individuals than among DM–non-HCC individuals (unadjusted OR 2.53, 95 % CI 1.66–3.87, $p = 1.05 \times 10^{-5}$). Of the other 50 SNPs, rs3785233 (which is located in an intron of *A2BP1*) and rs7754840 (which is located in an intron of *CDKAL1*) each showed a marginal association with the presence of HCC (unadjusted OR 1.63, 95 % CI 1.01–2.61, $p = 0.0422$ for *A2BP1*, unadjusted OR 1.54, 95 % CI 1.03–2.31, $p = 0.0336$ for *CDKAL1*).

The OR calculated for comparisons between the DM–HCC and DM-control groups with regard to *PNPLA3* rs738409 allele frequencies (unadjusted OR 3.22, 95 % CI 2.07–5.01, $p = 1.01 \times 10^{-7}$) were higher than those calculated for comparisons between the DM–HCC and DM–non-HCC groups. Likewise, the OR calculated for *A2BP1* rs3785233 allele frequencies in comparisons between the DM–HCC and DM-control groups (unadjusted OR 1.85, 95 % CI 1.11–3.08, $p = 0.0166$) were higher than those calculated for comparisons between the DM–HCC and DM–non-HCC groups.

The impact of JAZF1 on *HCC* development in patients homozygous for the PNPLA3 G allele with T2DM

A significant association of *PNPLA3* with HCC was observed not only with an allelic model but also with a recessive model (unadjusted OR 3.73, 95 % CI 2.02–6.88, $p = 1.33 \times 10^{-5}$) (Table 2). However, 43 patients homozygous for the *PNPLA3* G allele were in the DM-Control group; these individuals each had an LSM <7 kPa and had not developed HCC. To further clarify the genetic factors that might predispose individuals to develop HCC, we carried out association tests that only involved individuals homozygous for the *PNPLA3* G allele; specifically, we compared *PNPLA3* G homozygotes who developed HCC to those who did not (Supplementary Table 2). We found that the SNP rs864745, which is located in the intron of juxtaposed with another zinc finger gene 1 (*JAZF1*), showed a significant association with the presence of HCC both in a comparison between DM–HCC and DM–non-HCC *PNPLA3* G homozygotes (unadjusted OR 3.44, 95 % CI 1.77–6.71, $p = 0.0002$), and in a comparison between DM–HCC and DM-Control *PNPLA3* G homozygotes

Table 3 The SNPs associated with HCC among T2DM patients homozygous for the *PNPLA3* G allele

SNP No.	Gene	DM–HCC ($n = 30$)	DM–non-HCC ($n = 85$)	DM-control (n = 43)	DM–HCC vs. DM–non-HCC		DM–HCC vs. DM-control	
					Unadjusted OR [95 % CI]	*p* value	Unadjusted OR [95 % CI]	*p* value
Allelic model								
rs864745	*JAZF1* (G/A)	0.38/0.62	0.15/0.85	0.09/0.91	3.44 [1.77–6.71]	0.0002	6.06 [2.48–14.83]	2.44E−05
rs4523957	*SRR* (T/G)	0.9/0.1	0.72/0.28	0.67/0.33	3.44 [1.39–8.53]	0.0053	4.34 [1.67–11.31]	0.0015
rs391300	*SRR* (G/A)	0.9/0.1	0.74/0.26	0.69/0.31	3.24 [1.30–8.05]	0.0083	4.12 [1.58–10.74]	0.0024
rs574628	*ANGPT4* (G/A)	0.48/0.52	0.66/0.34	0.66/0.34	0.48 [0.27–0.88]	0.0164	0.48 [0.24–0.94]	0.0301
rs10811661	*CDKN2B* (T/C)	0.45/0.55	0.61/0.39	0.55/0.45	0.52 [0.29–0.94]	0.0296	0.68 [0.35–1.32]	0.2511
rs7903146	*TCF7L2* (T/C)	0/1	0.08/0.92	0.08/0.92	–	0.0218	–	0.0235
rs7901695	*TCF7L2* (T/C)	1/0	0.92/0.08	0.92/0.08	–	0.0218	–	0.0235
Recessive model								
rs864745	*JAZF1* (GG/nonGG)	0.23/0.77	0.01/0.99	0/1	25.57 [2.99–218.49]	4.11E−05	–	0.0009
rs4523957	*SRR* (TT/nonTT)	0.8/0.2	0.53/0.47	0.44/0.56	3.56 [1.32–9.58]	0.009298	5.05 [1.72–14.85]	0.0022
rs391300	*SRR* (GG/nonGG)	0.8/0.2	0.53/0.47	0.44/0.56	3.56 [1.32–9.59]	0.009298	5.05 [1.72–14.85]	0.0022
rs574628	*ANGPT4* (GG/nonGG)	0.23/0.77	0.41/0.59	0.37/0.63	0.43 [0.17–1.12]	0.080979	0.51 [0.18–1.46]	0.2092
rs10811661	*CDKN2B* (TT/nonTT)	0.17/0.83	0.36/0.64	0.26/0.74	0.35 [0.12–1.00]	0.044329	0.58 [0.18–1.89]	0.365
rs7903146	*TCF7L2* (TT/nonTT)	0/1	0/1	0/1	–	–	–	–
rs7901695	*TCF7L2* (TT/nonTT)	1/0	0.84/0.16	0.84/0.16	–	0.017695	–	0.0201

Table 4 Multivariate analysis of the effect of the *PNPLA3* G allele on HCC risk

Variables	DM–HCC vs. DM–non-HCC		DM–HCC vs. DM-control	
	Adjusted OR [95 % CI]	*p* value	Adjusted OR [95 % CI]	*p* value
PNPLA3 [G]	2.47 [1.60–3.82]	4.70E−05	2.83 [1.81–4.42]	4.73E−06
BMI	1.03 [0.95–1.13]	0.452933	1.09 [0.99–1.20]	0.0936
Sex (male)	2.16 [1.16–4.04]	0.015344	1.72 [0.89–3.32]	0.1066
Age	1.08 [1.04–1.13]	3.47E−05	1.06 [1.01–1.11]	0.0120

A multivariate logistic regression analysis was performed to adjust for any potential confounding effects (age, sex, BMI)

Table 5 Multivariate analysis of the effect of the *JAZF1* G allele on HCC risk among T2DM patients homozygous for the *PNPLA3* G allele

Variables	DM–HCC vs. DM–non-HCC		DM–HCC vs. DM-control	
	Adjusted OR [95 % CI]	*p* value	Adjusted OR [95 % CI]	*p* value
JAZF1 [G]	3.38 [1.64–6.95]	0.0009	4.38 [1.81–10.6]	0.0011
BMI	1.01 [0.88–1.17]	0.8400	1.10 [0.92–1.31]	0.2865
Sex (male)	2.51 [0.96–6.57]	0.0619	1.42 [0.48–4.20]	0.5226
Age	1.06 [1.00–1.12]	0.0339	1.02 [0.95–1.09]	0.6457

A multivariate logistic regression analysis was performed to adjust for any potential confounding effects (age, sex, BMI)

(unadjusted OR 6.06, 95 % CI 2.48–14.83, $p = 2.44 \times 10^{-5}$) (Table 3). Additionally, rs4523957 and rs391300, which are located within separate introns of serine racemase (*SRR*), each showed a marginal association with HCC in the comparison between the DM–HCC and DM-control groups (unadjusted OR 4.34, 95 % CI 1.67–11.31, $p = 0.0015$ for rs4523957, unadjusted OR 3.24, 95 % CI 1.30–8.05, $p = 0.0024$ for rs391300). These *SRR* SNPs, rs4523957 and rs391300, were in high linkage disequilibrium ($R^2 = 0.83$) with each other. The association of the *JAZF1* SNP and that of the *SRR* SNPs with HCC were stronger for the DM–HCC and DM-control comparisons than for the DM–HCC and DM–non-HCC comparisons.

Several other SNPs, which were located in *ANGPT4*, *CDKN2B*, or *TCF7L2*, showed minimal association with HCC in the comparison between the DM–HCC and DM–non-HCC groups.

Multivariate logistic regression analysis for *PNPLA3* and *JAZF1*

A multivariate logistic regression analysis was performed to identify any potential confounding non-genetic factors affecting HCC progression in patients with T2DM. Sex, age, and BMI were included along with the *PNPLA3* G allele in this analysis (Table 4). When comparing the DM–HCC and DM–non-HCC groups, three factors (*PNPLA3* G allele, sex, and age) were each identified as independent risk factors for HCC (*PNPLA3*: adjusted OR 2.47, 95 % CI 1.60–3.82, $p = 4.70 \times 10^{-5}$; sex: adjusted OR 2.16, 95 % CI 1.16–4.04, $p = 0.0153$; age: adjusted OR 1.08, 95 % CI 1.04–1.13, $p = 3.47 \times 10^{-5}$). Interestingly, in the comparison between the DM–HCC and DM–control groups, a stronger association between the *PNPLA3* G allele and HCC was observed (*PNPLA3*: adjusted OR 2.83, 95 % CI 1.81–4.42, $p = 4.73 \times 10^{-6}$).

An association of the *JAZF1* G allele with HCC in patients with T2DM and homozygous for the *PNPLA3* G allele was identified to be independent from three other potentially confounding factors (age, sex, and BMI) in both multivariate logistic regression analyses, the one comparing between the DM–HCC and DM–non-HCC groups (adjusted OR 3.38, 95 % CI 1.64–6.95, $p = 0.0009$) and the other comparing the DM–HCC and DM-control groups (adjusted OR 4.38, 95 % CI 1.81–10.60, $p = 0.0011$) (Table 5).

Even if the presence of cirrhosis was included as a factor in the multiple logistic regression analyses to adjust the influence of liver fibrosis, the *PNPLA3* G allele and the *JAZF1* G allele were independent risk factors for HCC in this cohort (Table 6). In addition, the multiple logistic regression analyses including platelet count instead of the presence of cirrhosis were performed, and it was revealed that the associations of *PNPLA3* and *JAZF1* with HCC were also independent from platelet count (Supplementary Tables 3, 4). These associations were also independent from the use of anti-diabetic medications (Supplementary Tables 5, 6).

Table 6 Multivariate analysis for *PNPLA3* or *JAZF1* including the presence of cirrhosis

Variables	DM–HCC vs. DM–non-HCC	
	Adjusted OR [95 % CI]	*p* value
(a) The effect of the PNPLA3 G allele on HCC risk		
PNPLA3 [G]	2.33 [1.42–3.84]	0.0009
BMI	1.04 [0.94–1.15]	0.493
Sex (male)	3.76 [1.72–8.22]	0.0009
Age	1.11 [1.06–1.17]	6.03E−06
Cirrhosis	22.32 [9.58–52.0]	6.12E−13
(b) The effect of the JAZF1 G allele on HCC risk among T2DM patients homozygous for the PNPLA3 G allele		
JAZF1 [G]	5.53 [2.11–14.52]	0.0005
BMI	1.02 [0.86–1.21]	0.8257
Sex (male)	7.15 [1.71–29.89]	0.007
Age	1.09 [1.01–1.18]	0.0296
Cirrhosis	45.32 [8.94–229.72]	4.12E−06

Discussion

In 2007, the Japan Society of DM reported that malignancies were the most frequent cause of death (34.1 %) among 18,385 patients with DM during 1991–2000, and that among the malignancies, HCC showed the highest frequency (8.6 %). Surprisingly, the frequency of deaths caused by liver cirrhosis was 4.7 %; cumulatively, 13.3 % of these patients with DM died of liver diseases; notably, neither the incidences of HBV or HCV infection, nor the quantity of alcohol intake were reported in the study [24]. Subsequently, Shima et al. concluded that the majority of liver injuries in Japanese DM patients were associated with NAFLD/NASH [21], and currently, NASH is considered to be a leading cause of non-HBV and non-HCV- related HCC.

In 2008, Romeo reported an association between the risk of liver fat accumulation and the *PNPLA3* SNP rs738409 [19]. Subsequently, it was reported that the *PNPLA3* variant was significantly associated with the severity of NAFLD [25, 26] and alanine aminotransferase (ALT) levels [19, 27]; additionally, the association of *PNPLA3* was validated in different populations around the world, from children to adults [28, 29]. In addition, this *PNPLA3* SNP is associated with HCC in European Caucasian patients with NAFLD [20, 30], but an association between *PNPLA3* and HCC was not replicated in a study of Japanese patients with NAFLD [31].

Here, we found that the rs738409 C >G polymorphism was significantly associated with HCC, even in our restricted cohort of Japanese patients with T2DM. This effect was independent of potentially confounding factors including age, sex, BMI, and even the presence of cirrhosis. We believe that NASH may have been a confounding factor in our study because T2DM was present in 59 % of patients with NASH who developed HCC and had participated in a cross-sectional multicenter study in Japan [32]. However, only 24.5 % (12/49) of DM-HCC patients who had given a liver biopsy specimen were diagnosed with steatohepatitis based on histology in our current study. Thus, larger prospective cohort studies should be conducted for further examination of associations of *PNPLA3* with HCC in Japanese patients with NAFLD/NASH.

A previously unidentified susceptibly SNP for HCC was identified by stratifying our cohort by *PNPLA3* genotype. The analysis involving only the group of individuals with the *PNPLA3* GG genotype identified that the SNP rs864745 located in *JAZF1* increases susceptibility to HCC among patients with T2DM and homozygous for the *PNPLA3* G allele. The frequency of the *JAZF1* rs864745 G allele was significantly higher among DM-HCC individuals than among DM-non HCC or DM-Control individuals, and this association was independent of age, sex, BMI, and even the presence of cirrhosis.

The SNP rs864745 resides within intron 1 of *JAZF1*, which encodes a transcriptional repressor [33]. *JAZF1* is expressed in the pancreas, and expression of *JAZF1* is downregulated in patients with T2DM [34]. Notably, *JAZF1* rs864745 variants influence traits of insulin secretion [33, 35]. According to these reports, *JAZF1* rs864745 is associated with susceptibility to T2DM because it is associated with a lower acute insulin response, and insulin is considered to be a background factor that influences the onset and progression of cancer [36, 37].

Here we found that the *JAZF1* rs864745 G allele is a risk factor for development of HCC. The studies that identified the association of this SNP with T2DM also show that the A allele is associated with T2DM. Although our hypothesis

was that risk alleles for HCC would be the same as those for T2DM, our results showed that the JAZF1 A allele, which confers susceptibility to T2DM, reduced the risk of HCC development in patients with the *PNPLA3* GG genotype who had T2DM. The SNP rs864745 in *JAZF1* also influences susceptibility to colorectal cancer [6]. Interestingly, the A allele of rs864745, which is associated with increased risk for T2DM, is also associated with a decreased risk of colorectal cancer; this pattern is similar to the pattern of association between *JAZF1* and HCC found in this study. *JAZF1* also reportedly influences susceptibility to prostate cancer [38].

Several additional T2DM susceptibility loci showed a trend towards significant association with HCC. Associations between variants located in *A2BP1* (rs3785233) or *SRR* (rs4523957 and rs391300) and HCC were convincing when DM-HCC patients were compared with DM-Control patients instead of DM-non-HCC patients. A large-scale clinical study is required to confirm or refute these associations because they may be potential candidates of low-penetrance genes for susceptibility to HCC in Japanese patients with T2DM.

The current study provides the proof-of-concept for our approach to searching for susceptible genes influencing HCC development; specifically, the list of known candidate genes involved in the pathogenesis of comorbid T2DM is probably highly enriched with HCC in T2DM susceptibility genes. A similar approach could be useful for analyses of other diseases; especially when the incidence of certain cancers is higher in patients of which comorbidity is predisposed to genetic influence.

In conclusion, the *PNPLA3* rs738409 was determined to be associated with HCC development in a cohort of Japanese patients with T2DM. Additionally, the *JAZF1* rs864745 was firstly identified as a risk factor for HCC among T2DM patients with the GG genotype at *PNPLA3* rs738409. We believe that inclusion of these SNPs into multi-factorial risk assessments may help physicians to identify T2DM patients who have a high risk of developing HCC among the expanding population of T2DM patents.

Acknowledgments We thank Ms. Yoriko Mawatari, Ms. Mayumi Ishii, Ms. Takayo Tsuchiura, Ms. Kozue Sugimoto, and Ms. Noriko Ota for their technical support. This study was supported by grants (25-202, 26-206) from the National Center for Global Health and Medicine in Japan and the Ministry of Education, Culture, Sports, Science, and Technology (No. 25461019) and by a research award from the Liver Forum in Kyoto; it was also supported in part by the Ministry of Health, Labour, and Welfare of Japan.

Compliance with ethical standards

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

References

1. International Diabetes Federation. IDF Diabetes Atlas, 6th edn. Brussels, Belgium: International Diabetes Federation, 2013. http://www.idf.org/diabetesatlas.
2. Sasazuki S, Charvat H, Hara A, et al. Diabetes mellitus and cancer risk: pooled analysis of eight cohort studies in Japan. Cancer Sci. 2013;104:1499–507.
3. Seshasai SR, Kaptoge S, Thompson A, et al. Diabetes mellitus, fasting glucose, and risk of cause-specific death. N Engl J Med. 2011;364:829–41.
4. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology. 2007;132:2557–76.
5. Frayling TM, Colhoun H, Florez JC. A genetic link between type 2 diabetes and prostate cancer. Diabetologia. 2008;51:1757–60.
6. Cheng I, Caberto CP, Lum-Jones A, et al. Type 2 diabetes risk variants and colorectal cancer risk: the multiethnic cohort and page studies. Gut. 2011;60:1703–11.
7. Miyake K, Yang W, Hara K, et al. Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong evidence of the association. J Hum Genet. 2009;54:236–41.
8. Lyssenko V, Jonsson A, Almgren P, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med. 2008;359:2220–32.
9. Talmud PJ, Hingorani AD, Cooper JA, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. BMJ. 2010;340:b4838.
10. Omori S, Tanaka Y, Takahashi A, et al. Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8, and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. Diabetes. 2008;57:791–5.
11. Yasuda K, Miyake K, Horikawa Y, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet. 2008;40:1092–7.
12. Tsai FJ, Yang CF, Chen CC, et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet. 2010;6:e1000847.
13. Maeda S, Kobayashi MA, Araki S, et al. A single-nucleotide polymorphism within the acetyl-coenzyme A carboxylase beta gene is associated with proteinuria in patients with type 2 diabetes. PLoS Genet. 2010;6:e1000842.
14. Okamoto K, Iwasaki N, Nishimura C, et al. Identification of KCNJ15 as a susceptibility gene in Asian patients with type 2 diabetes mellitus. Am J Hum Genet. 2010;86:54–64.
15. Leak TS, Langefeld CD, Keene KL, et al. Chromosome 7p linkage and association study for diabetes related traits and type 2 diabetes in an African–American population enriched for nephropathy. BMC Med Genet. 2010;11:22.
16. Tanaka N, Babazono T, Saito S, et al. Association of solute carrier family 12 (sodium/chloride) member 3 with diabetic nephropathy, identified by genome-wide analyses of single-nucleotide polymorphisms. Diabetes. 2003;52:2848–53.
17. Kamiyama M, Kobayashi M, Araki S, et al. Polymorphisms in the 3′ UTR in the neurocalcin delta gene affect mRNA stability, and confer susceptibility to diabetic nephropathy. Hum Genet. 2007;122:397–407.
18. Shimazaki A, Kawamura Y, Kanazawa A, et al. Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. Diabetes. 2005;54:1171–8.
19. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2008;40:1461–5.

20. Liu YL, Patman GL, Leathart JB, et al. Carriage of the PNPLA3 rs738409 C > G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma. J Hepatol. 2014;61:75–81.
21. Shima T, Uto H, Ueki K, et al. Clinicopathological features of liver injury in patients with type 2 diabetes mellitus and comparative study of histologically proven nonalcoholic fatty liver diseases with or without type 2 diabetes mellitus. J Gastroenterol. 2013;48:515–25.
22. Gaia S, Carenzi S, Barilli AL, et al. Reliability of transient elastography for the detection of fibrosis in non-alcoholic fatty liver disease and chronic viral hepatitis. J Hepatol. 2011;54:64–71.
23. Nishida N, Tanabe T, Takasu M, et al. Further development of multiplex single-nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem. 2007;364:78–85.
24. Hotta N, Nakamura J, Iwamoto Y, et al. Causes of death in Japanese diabetics: a questionnaire survey of 18,385 diabetics over a 10-year period. J Diabetes Investig. 2010;1:66–76.
25. Kawaguchi T, Sumida Y, Umemura A, et al. Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. PLoS One. 2012;7:e38322.
26. Sookoian S, Castaño GO, Burgueño AL, et al. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res. 2009;50:2111–6.
27. Yuan X, Waterworth D, Perry JR, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet. 2008;83:520–8.
28. Sookoian S, Pirola CJ. Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. Hepatology. 2011;53:1883–94.
29. Sookoian S, Pirola CJ. PNPLA3, the history of an orphan gene of the potato tuber protein family that found an organ: the liver. Hepatology. 2014;59:2068–71.
30. Trépo E, Nahon P, Bontempi G, et al. Association between the PNPLA3 (rs738409 C > G) variant and hepatocellular carcinoma: evidence from a meta-analysis of individual participant data. Hepatology. 2014;59:2170–7.
31. Takeuchi Y, Ikeda F, Moritou Y, et al. The impact of patatin-like phospholipase domain-containing protein 3 polymorphism on hepatocellular carcinoma prognosis. J Gastroenterol. 2013;48:405–12.
32. Yasui K, Hashimoto E, Komorizono Y, et al. Characteristics of patients with nonalcoholic steatohepatitis who develop hepatocellular carcinoma. Clin Gastroenterol Hepatol. 2011;9:428–33 **(quiz e450)**.
33. Grarup N, Andersen G, Krarup NT, et al. Association testing of novel type 2 diabetes risk alleles in the JAZF1, CDC123/CAMK1D, TSPAN8, THADA, ADAMTS9, and NOTCH2 loci with insulin release, insulin sensitivity, and obesity in a population-based sample of 4,516 glucose-tolerant middle-aged Danes. Diabetes. 2008;57:2534–40.
34. Marselli L, Thorne J, Dahiya S, et al. Gene expression profiles of Beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. PLoS One. 2010;5:e11499.
35. Gjesing AP, Hornbak M, Allin KH, et al. High heritability and genetic correlation of intravenous glucose- and tolbutamide-induced insulin secretion among non-diabetic family members of type 2 diabetic patients. Diabetologia. 2014;57:1173–81.
36. Kawaguchi T, Izumi N, Charlton MR, et al. Branched-chain amino acids as pharmacological nutrients in chronic liver disease. Hepatology. 2011;54:1063–70.
37. Kasuga M, Ueki K, Tajima N, et al. Report of the Japan Diabetes Society/Japanese Cancer Association Joint Committee on diabetes and cancer. Cancer Sci. 2013;104:965–76.
38. Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet. 2008;40:310–5.