GWAS-Identified Common Variants With Nonalcoholic Fatty Liver Disease in Chinese Children

Xiao-Rui Shang, Jie-Yun Song, Fang-Hong Liu, Jun Ma, and Hai-Jun Wang

ABSTRACT

Objectives: Three genome-wide association studies were previously done for nonalcoholic fatty liver disease (NAFLD) among individuals of Western countries and identified several genetic variants associated with NAFLD. The study aimed to identify whether 7 GWAS-identified common variants (*GCKR* rs780094, *PDGFA* rs343064, *FDFT1* rs2645424, *COL13A1* rs1227756, *EHBP1L1* rs6591182, *NCAN* rs2228603, and *PNPLA3* rs738409) were associated with NAFLD in Chinese children.

Methods: This case-control study recruited 1027 Chinese children of age 7 to 18 years, including 162 children with NAFLD and 865 children without NAFLD. Anthropometric measurements, alanine transaminase (ALT) detection, liver ultrasound examination, and genotyping of 7 variants were performed.

Results: The G-allele of *PNPLA3* rs738409 was associated with NAFLD (odds ratio [OR] 1.55, 95% confidence interval 1.13–2.11, $P = 0.006$) and moderate-to-severe steatosis (OR 3.77, 95% confidence interval 1.78–7.98, $P = 0.001$) adjusted for age, sex, and BMI standard deviation score. In addition, we found each G-allele of rs738409 increased ALT level by 1.09 IU/L ($P = 0.011$). Subjects carrying 10 or more risk alleles of 7 variants had an OR of 4.76 ($P = 0.025$) for NAFLD compared with subjects carrying 3 or fewer risk alleles.

Conclusions: The *PNPLA3* rs738409 G-allele was associated with NAFLD and ALT level in Chinese children. It had stronger association with moderate-to-severe steatosis. Children carrying 10 or more risk alleles of 7 variants were susceptible for NAFLD.

Key Words: children, gene, nonalcoholic fatty liver disease, polymorphism

(*JPGN* 2015;60: 669–674)

What Is Known

- Genetic factors play an important role in nonalcoholic fatty liver disease (NAFLD) susceptibility.
- Three recent genome-wide association studies among individuals from Western countries have identified multiple genetic variants that are associated with NAFLD.
- Most genome-wide association studies–identified single nucleotide polymorphisms were not studied for NAFLD in a Chinese population.

What Is New

- The *PNPLA3* rs738409 was associated with NAFLD and alanine transaminase level and had stronger association with moderate-to-severe steatosis in Chinese children.
- The association of rs738409 with NAFLD in obese children was stronger than that in nonobese children.
- Children carrying 10 or more risk alleles of 7 variants were susceptible for NAFLD.

Nonalcoholic fatty liver disease (NAFLD) is a common chronic liver disease, including a spectrum of disease ranging from fatty infiltration of the liver (steatosis) to histologic evidence of inflammation (nonalcoholic steatohepatitis), to fibrosis or cirrhosis, without a history of excessive alcohol ingestion (1). It affects 20% to 30% of the population in Western countries (2). Although the prevalence of NAFLD is somewhat lower in east Asia, that is 18% in South Korea (3) and 15% in China (4), its prevalence has increased rapidly in young generations during the last 2 decades (5). Studies have shown that the prevalence of NAFLD in children is approximately 2.6% to 9.6%, and it reaches up to 32.3% to 43.9% in obese children (6). Pediatric NAFLD is a growing global health problem worldwide (7). Children with NAFLD may develop end-stage liver disease with a consequent need for liver transplantation (8).

Genetic underpinning of NAFLD has been supported by familial aggregation studies (9,10), heritability studies (11,12), candidate gene studies (13), genome-wide association studies (GWAS) (14–16), and gene expression studies (17,18). Several candidate genes have been implicated in the pathogenesis of NAFLD, including genes involved in hepatic lipid metabolism, insulin sensitivity, generation of reactive oxidant species, or cytokine (19). The results from candidate gene studies have, however, been rather disappointing because of conflicting results and the absence of replication in independent populations. Recent advances in genotyping technology together with detailed characterization of

Received February 14, 2014; accepted December 3, 2014.

From the Institute of Child and Adolescent Health, School of Public Health, Peking University, Beijing, China.

Address correspondence and reprint requests to Dr Hai-Jun Wang, Professor, Institute of Child and Adolescent Health, School of Public Health, Peking University, No. 38 Xueyuan Road, Haidian District, Beijing 100-191, China (e-mail: whjun1@bjmu.edu.cn).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (*www.jpgn.org*).

The study was supported by grants from the Major State Basic Research and Development Program of China (973 program) (2012CB517501) and National Natural Science Foundation of China (81172683).

The authors report no conflicts of interests.

DOI: 10.1097/MPG.0000000000000662

common gene variants have led to a rapid development of GWAS. Until the end of 2012, 3 GWAS for NAFLD among individuals of Western countries have been published for NAFLD (14–16).

The major problem in both candidate gene studies and GWAS is the presence of false-negative and false-positive associations. Consequently, many journals extremely recommend authors to replicate their findings in different samples before accepting the new single nucleotide polymorphism (SNP) as truly associated with NAFLD (19). So we selected 7 SNPs identified by GWAS with the minor allele frequency >0.10 in Chinese populations (rs780094 in the glucokinase (hexokinase 4) regulator gene (*GCKR*), rs343064 in the platelet-derived growth factor alpha polypeptide gene (*PDGFA*), rs2645424 in the farnesyl-diphosphate farnesyltransferase 1 gene (*FDFT1*), rs1227756 in the collagen, type XIII, alpha 1 gene (*COL13A1*), rs6591182 in the EH domain binding protein 1-like 1 gene (*EHBP1L1*), rs2228603 in the neurocan gene (*NCAN*), and rs738409 in the patatin-like phospholipase domain containing 3 gene (*PNPLA3*)), to analyze their relation with NAFLD in Chinese children. Among these 7 SNPs, the association of *GCKR* rs780094 and *PNPLA3* rs738409 with NAFLD had been studied in Chinese populations (20–26). The minor allele frequency of either rs780094 (0.47) or rs738409 (0.37) in Taiwanese children (25,26) was similar to that of Europeans and Americans in previous GWAS studies (0.41 for rs780094, 0.17–0.49 for rs738409). In addition, the effects of these 2 SNPs reported in previous studies were not different between Chinese and other ethnic populations. Until now, the other 5 SNPs have not studied for NAFLD among Chinese population. Investigating these SNPs in different populations would be helpful to evaluate the generalizability of these findings and identify causal variants for NAFLD. The present study aimed to determine whether the 7 common variants are associated with NAFLD and alanine transaminase (ALT) level in Chinese children.

METHODS

Subjects

Subjects were 1093 individuals ages from 7 to 18 years who participated in the Comprehensive Prevention Project for Overweight and Obese Adolescents in Beijing, China. They were recruited from 3 middle schools and 2 elementary schools of the Haidian District of Beijing. The ascertainment strategy for the study population has been described in detail previously (27). By asking medical history, we selected the subjects without any of the following conditions: alcohol consumption; a history of diseases or drugs (including herbal medicines) causing liver disease; common (hepatitis B virus, hepatitis C virus) or rare liver diseases; hepatic malignancies; infections biliary tract disease; and any cardiovascular and metabolic diseases. Finally, 1027 children having liver ultrasound examination and blood samples were included in the study. All of the participants gave their written informed consent. The study was approved by the ethic committee of Peking University Health Science Center.

Measurements

Anthropometric measurements, including height, weight, waist circumference, systolic blood pressure (SBP), and diastolic blood pressure (DBP), were performed according to standard protocols. Fasting venous blood samples were taken for detection of ALT, triglyceride (TG), total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and fasting glucose (FPG) by using a biochemical auto-analyzer (Hitachi 7060, Tokyo, Japan). Liver ultrasound examination was performed by 1 experienced doctor with an ultrasound system (SIEMENS Soneline G50, Berlin, Germany). The sex- and age-specific body mass index standard deviation score (BMI-SDS) was calculated by using the growth reference data of the World Health Organization for children and adolescents of age 5 to 19 years (*www.whoint/childgrowth/standards/bmi_for_age/en/indexhtml*).

The subjects were confirmed to have hepatic steatosis by liver ultrasonography and classified into 3 categories—mild, moderate, and severe steatosis—according to reference criteria (28). The criteria included the following: diffuse enhancement of near field echo in the hepatic region and gradual attenuation of the far field echo; unclear display of intrahepatic lacuna structure; mild-to-moderate hepatomegaly with a round blunt border; reduction of blood flow signal in the liver; and unclear or nonintact display of envelop of right liver lobe and diaphragm. Patients meeting criterion 1 and any 1 of criteria 2 to 4 were classified as mild; patients meeting criterion 1 and any 2 of criteria 2 to 4 were classified as moderate; and patients meeting criteria 1, 5 and any 2 of criteria 2 to 4 were classified as severe. A previous study provided data that the staging system used in the criteria correlates well to histology (29). All of the examinations were performed by 1 experienced doctor, who was unaware of the patients' clinical details and laboratory findings.

We used the BMI percentile criteria for obese and nonobese children, which were determined in a representative Chinese population (30). According to the criteria, the children with an age- and sex-specific BMI greater than or equal to the 95th percentile are defined as obese, whereas those with a BMI between 15th and 95th percentile are nonobese. We calculated the prevalence of high SBP or DBP in different groups, based on the age- and sex-specific blood pressure criteria for Chinese children (31).

Selection of SNPs and Genotyping

Until the end of 2012, 11 SNPs were reported to be associated with NAFLD by 3 GWAS previously done among individuals of Western countries (14–16). Considering statistical power, from these 11 SNPs we selected 7 SNPs (*GCKR* rs780094, *PDGFA* rs343064, *FDFT1* rs2645424, *COL13A1* rs1227756, *EHBP1L1* rs6591182, *NCAN* rs2228603, and *PNPLA3* rs738409), with the minor allele frequency >0.10 in Chinese populations (Hapmap database: *http://www.hapmap.org*). With the assumed effect size (odds ratio [OR] = 1.5) and effect allele frequency ≥0.10, statistical power to detect a positive association would be >0.70, given our sample size.

Genomic DNAs of subjects were extracted from blood leukocytes by the phenol–chloroform extraction method. Genotyping was conducted on MassARRAY System (Sequenom, San Diego, CA). Primers, including a pair of amplification primers and an extension primer for each SNP, were designed with Sequenom MassArray Assay Design Suite. A multiplex polymerase chain reaction was performed, and unincorporated double-stranded nucleotide triphosphate bases were dephosphorylated with shrimp alkaline phosphatase followed by primer extension. The purified primer extension reaction was spotted on to a 384-element silicon chip (SpectroCHIP; Sequenom) and analyzed in the matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS; Sequenom). The resulting spectra were processed with MassArray Typer (Sequenom) (*www.sequenom.com*). The genotyping call rate of rs2645424, rs6591182, and rs780094 was 99.9%, and the other 4 SNPs had a call rate of 100%. All of the experiments were done by investigators who were blind to the phenotypes.

Statistical Analyses

Statistical analyses were performed using SPSS 18.0 software (IBM SPSS, Armonk, NY) and PLINK (Massachusetts

General Hospital, Boston, MA) (32). Differences in general characteristics between the NAFLD groups and the control group were tested with *t* test (continuous variables) or χ^2 test (category variables). The genotype data of control group were tested for deviation from Hardy-Weinberg equilibrium. F-statistics (F_{ST}), a metric representation of the effect of population subdivision, was calculated according to the following formula, $F_{ST} = (P_1 - P_2)^2 / ((P_1 + P_2) \times (2 - (P_1 + P_2)))$, where P_1 = allele frequency in the population of previous GWAS study and P_2 = allele frequency in the population of our study (33,34). A F_{ST} value between 0 and 0.05 indicates little genetic differentiation; a value between 0.05 and 0.15, moderate differentiation; a value between 0.15 and 0.25, great differentiation; and values >0.25, very great differentiation (35).

Multivariate logistic regression with age, sex, and BMI-SDS as covariates was used to calculate the ORs of the genetic variants for NAFLD. The relations between the variants and ALT level or other metabolic phenotypes (including waist circumference, blood pressure, serum lipids, and FPG) were tested by using the linear regression analysis adjusted for age, sex, and BMI-SDS. We created a genetic risk score (GRS) by summing the number of NAFLD-susceptible alleles of the 7 variants. The NAFLD-susceptible alleles were determined based on the literatures of GWAS (14–16). Logistic regression was used to examine the OR of the GRS score for NAFLD. All of the variants were analyzed under the additive model based on recent publications (26,36). A 2-sided *P* value <0.05 was considered as nominal significant. Adjustment was made for multiple testing using Bonferroni correction.

RESULTS

The general characteristics of patients with NAFLD and controls are shown in Table 1. The NAFLD group consisted of 162 children (47 girls, mean age 11.81 ± 2.20 years, mean BMI = 26.75 ± 3.85 kg/m^2), and the control group consisted of 865 children (406 girls, mean age 11.44 ± 2.99 years, mean BMI = 20.72 ± 3.61 kg/m^2). Sex differed between 2 groups ($P < 0.001$), whereas age difference between the 2 groups was not statistically significant ($P = 0.112$). Compared with controls, the NAFLD group had lower high-density lipoprotein cholesterol and higher BMI, BMI-SDS, waist circumference, ALT, total cholesterol, TG, low-density lipoprotein cholesterol, SBP, DBP, and FPG ($P < 0.05$). Overall 29.2% of the children were obese. There were more obese children in the NAFLD group compared with the control group (74.7% vs 20.2%, $P < 0.001$). The prevalence of high SBP or DBP was higher in the NAFLD group compared with controls. The characteristics of subjects with mild NAFLD and moderate-to-severe steatosis are also reported in Table 1.

The genotyping information of 7 SNPs and genotype distributions in subjects with and without NAFLD are shown in supplementary Table 1 (*http://links.lww.com/MPG/A412*). The genotype distribution of all of the variants in the control group was in Hardy–Weinberg equilibrium ($P > 0.05$). Based on the F_{ST} values, we identified 2 SNPs (*FDFT1* rs2645424 and *COL13A1* rs1227756) having moderate differentiation (F_{ST} between 0.05 and 0.15) between our Chinese population and the Western populations in previous GWAS studies, whereas 5 SNPs have little differentiation ($F_{ST} < 0.05$).

As shown in Table 2, in the multivariate logistic regression analyses with age, sex, and BMI-SDS as covariates, the rs738409 G-allele was strongly associated with NAFLD (OR 1.55, 95% confidence interval [CI] 1.13–2.11, $P = 0.006$). Subsequently, we analyzed the relation between the SNPs and moderate-to-severe steatosis. The rs738409 was significantly associated with moderate-to-severe steatosis under the additive model adjusted for age, sex, and BMI-SDS (OR 3.77, 95% CI 1.78–7.98, $P = 0.001$). The association of rs738409 with NAFLD or moderate-to-severe steatosis was still significant after Bonferroni correction for multiple testing ($P < 0.007$, 0.05 divided by 7). No significant association was found between other SNPs and NAFLD.

As shown in Table 3, when we stratified the subjects into nonobese and obese groups, we observed the association of rs738409 G-allele with NAFLD in obese children (OR 1.85, 95% CI 1.22–2.81, $P = 0.004$) was stronger than that in nonobese children (OR 1.17, 95% CI 0.71–1.92, $P = 0.541$).

Using linear regression adjusted for age, sex, and BMI-SDS, we found significant association between *PNPLA3* rs738409 and ALT level (Table 4). Each G-allele of rs738409 increased ALT

TABLE 1. General characteristics of patients with NAFLD and controls

	NAFLD cases			
	Mild steatosis (n = 132)	Moderate-to-severe steatosis (n = 30)	Total NAFLD cases (n = 162)	Controls (n = 865)
Female/male	40/92	7/23	47/115*	406/459
Age	11.64 ± 2.23	12.60 ± 1.92	11.81 ± 2.20	11.44 ± 2.99
BMI, kg/m^2	26.05 ± 3.24	29.82 ± 4.79	26.75 ± 3.85*	20.72 ± 3.61
BMI-SDS	2.38 ± 0.72	2.74 ± 0.69	2.44 ± 0.72*	0.92 ± 1.16
Waist circumference, cm	83.96 ± 9.96	94.75 ± 10.97	85.97 ± 10.96*	69.32 ± 10.03
ALT, IU/L	20.31 ± 16.00	36.92 ± 22.11	23.04 ± 18.16*	12.29 ± 7.05
TG, mmol/L	1.09 ± 0.48	1.54 ± 0.70	1.17 ± 0.56*	0.87 ± 0.36
TC, mmol/L	4.22 ± 0.72	4.25 ± 0.90	4.23 ± 0.75*	4.09 ± 0.67
LDL-C, mmol/L	2.47 ± 0.61	2.62 ± 0.71	2.50 ± 0.63*	2.17 ± 0.54
HDL-C, mmol/L	1.40 ± 0.27	1.20 ± 0.25	1.36 ± 0.28*	1.54 ± 0.31
SBP, mmHg	112.76 ± 14.87	124.20 ± 16.34	114.88 ± 15.74*	104.45 ± 13.85
Abnormal SBP	34 (25.8%)	13 (43.3%)	47 (29.0%)	87 (10.1%)
DBP, mmHg	66.95 ± 10.09	71.53 ± 9.72	67.80 ± 10.15*	62.36 ± 9.32
Abnormal DBP	14 (10.6%)	7 (23.3%)	21 (13.0%)	21 (2.4%)
FPG, mmol/L	5.48 ± 0.44	5.49 ± 0.39	5.48 ± 0.43*	5.34 ± 0.40

ALT = alanine aminotransferase; BMI-SDS = body mass index–standard deviation score; DBP = diastolic blood pressure; FPG = fasting glucose; HDL-C = high-density lipoprotein cholesterol; LDL-C = low-density lipoprotein cholesterol; NAFLD = nonalcoholic fatty liver disease; SBP = systolic blood pressure; SD = standard deviation; TC = total cholesterol; TG = triglyceride. Results were presented as mean ± SD.

* $P < 0.05$ when compared with the control group.

TABLE 2. Association of 7 SNPs with NAFLD in Chinese children

SNP	Nearest gene	Effect allele	NAFLD vs control		Moderate-to-severe steatosis vs control	
			OR (95% CI)	Nominal *P*	OR (95% CI)	Nominal *P*
rs780094	*GCKR*	A	1.19 (0.89–1.61)	0.234	0.74 (0.37–1.47)	0.383
rs343064	*PDGFA*	T	1.23 (0.91–1.64)	0.175	0.59 (0.31–1.15)	0.121
rs2645424	*FDFT1*	C	0.97 (0.69–1.36)	0.859	0.79 (0.36–1.74)	0.563
rs1227756	*COL13A1*	G	1.08 (0.77–1.51)	0.66	1.70 (0.73–3.96)	0.217
rs6591182	*EHBP1L1*	A	0.98 (0.73–1.30)	0.867	1.74 (0.88–3.41)	0.109
rs2228603	*NCAN*	T	0.97 (0.59–1.59)	0.894	0.45 (0.12–1.63)	0.223
rs738409	*PNPLA3*	G	1.55 (1.13–2.11)	0.006	3.77 (1.78–7.98)	0.001

Logistic analyses were conducted under the additive model with age, sex, and BMI-SDS as covariates. Moderate-to-severe steatosis was classified according to reference criteria (29). BMI-SDS = body mass index–standard deviation score; CI = confidence interval; COL13A1 = collagen, type XIII, alpha 1; EHBP1L1 = EH domain binding protein 1-like 1 gene; FDFT1 = farnesyl-diphosphate farnesyltransferase 1; GCKR = glucokinase (hexokinase 4) regulator; NAFLD = nonalcoholic fatty liver disease; NCAN = neurocan; OR = odds ratio; PDGFA = platelet-derived growth factor alpha polypeptide; PNPLA3 = patatin-like phospholipase domain containing 3; SNP = single nucleotide polymorphism.

level by 1.09 IU/L (95% CI 0.25–1.93, $P = 0.011$). The association did not exist after Bonferroni correction.

We also analyzed the association between these 7 SNPs and other metabolic phenotypes (including waist circumference, blood pressure, serum lipids, and FPG). As shown in supplementary Tables 2 and 3 (*http://links.lww.com/MPG/A413* and *http://links.lww.com/MPG/A414*, we only identified the association between *GCKR* rs780094 and TG, which was significant after Bonferroni correction for multiple testing ($\beta = 0.08$, 95% CI 0.05–0.11, $P = 2.53 \times 10^{-6} < 0.0009$, 0.05 divided by 7 SNPs and 8 phenotypes).

To estimate the cumulative effect of the 7 SNPs on NAFLD, we constructed the GRS by summing the number of NAFLD-susceptible alleles carried by each individual. On average, each additional effect allele was associated with 13% increased risk of NAFLD (OR 1.13 95% CI 1.00–1.28, $P = 0.046$). Subjects carrying 10 or more risk alleles had an OR of 4.76 (95% CI 1.22–18.59, $P = 0.025$) for NAFLD compared with subjects that carry 3 or fewer risk alleles (Fig. 1).

DISCUSSION

In this study, we tested association of NAFLD with 7 common variants that were identified by 3 GWAS. Our study revealed that a nonsynonymous variant rs738409 isoleucine-to-methionine change at the amino acid 148 in *PNPLA3* was associated with NAFLD, and had stronger association with moderate-to-severe steatosis.

In the present study, we found that the G-allele of *PNPLA3* rs738409 was associated not only with NAFLD (OR 1.55, 95% CI 1.13–2.11, $P = 0.006$) but also with moderate-to-severe steatosis having a higher OR (OR 3.77, 95% CI 1.78–7.98, $P = 0.001$) in Chinese children. PNPLA3 is predominantly expressed in human liver and adipose tissue, possesses both lipolytic and lipogenic activity in vitro, and localizes to the surface of lipid droplets in hepatocytes. The rs738409 represents a cytosine to guanine substitution, resulting in decreased enzymatic activity and hepatocellular fat accumulation (37). Recent studies have shown that the rs738409 variant in the *PNPLA3* leads to hepatic steatosis and steatohepatitis by enhancing the lipogenic activity and impairing the lipolytic activity of the PNPLA3 in mouse liver (38). The rs738409 in *PNPLA3* gene was identified in 2 GWAS as the strongest genetic determinant of hepatic steatosis and increased ALT levels. Since then, the association was replicated in several independent candidate gene studies (39) and a meta-analysis across different populations (40). Our results indicated that the SNP had a strong effect on liver fat accumulation in Chinese children, which validate the effects of the SNP on NAFLD in children of different ethnicity.

By stratified analyses, we found the association of rs738409 G-allele with NAFLD in obese children (OR 1.85, 95% CI 1.22–2.81, $P = 0.004$) was stronger than that in nonobese children (OR 1.17, 95% CI 0.71–1.92, $P = 0.541$). Our results indicated that the presence of the G-allele can increase the effect of obesity on the risk of developing NAFLD in children, suggesting combined effect of rs738409 and obesity. The similar effect of the rs738409 and BMI on NAFLD in adults was previously reported by Peng et al (21). They suggested that rs738409 polymorphism had an additive effect with obesity in increasing NAFLD risk (interaction contrast ratio 2.31, 95% CI 0.70–8.86). There was a study showing interaction

TABLE 3. Association of the rs738409 with NAFLD in nonobese and obese groups

Group	GG n (%)	GC n (%)	CC n (%)	OR (95% CI)	*P*
Nonobese					
Controls	94 (13.6)	337 (48.8)	259 (37.5)		
NAFLD cases	8 (19.5)	19 (46.3)	14 (34.1)	1.17 (0.71–1.92)	0.541
Obese					
Controls	15 (8.6)	81 (46.3)	79 (45.1)		
NAFLD cases	20 (16.5)	55 (45.5)	46 (38.0)	1.85 (1.22–2.81)	0.004

BMI-SDS = body mass index–standard deviation score; CI = confidence interval; NAFLD = nonalcoholic fatty liver disease; OR = odds ratio. Logistic analyses were conducted under the additive model with age, sex, and BMI-SDS as covariates.

TABLE 4. Association of 7 SNPs with ALT level in Chinese children

SNP	Nearest gene	Effect allele	b	SEM	Nominal *P*
rs780094	*GCKR*	A	0.74	0.41	0.071
rs343064	*PDGFA*	T	−0.62	0.40	0.123
rs2645424	*FDFT1*	C	−0.43	0.47	0.369
rs1227756	*COL13A1*	G	0.44	0.45	0.325
rs6591182	*EHBP1L1*	A	0.33	0.40	0.420
rs2228603	*NCAN*	T	1.00	0.68	0.141
rs738409	*PNPLA3*	G	1.09	0.43	0.011

Linear regressions were analyzed under the additive model with age, sex, and BMI-SDS as covariates. ALT = alanine aminotransferase; BMI-SDS = body mass index–standard deviation score; COL13A1 = collagen, type XIII, alpha 1; EHBP1L1 = EH domain binding protein 1-like 1 gene; FDFT1 = farnesyl-diphosphate farnesyltransferase 1; GCKR = glucokinase (hexokinase 4) regulator; NCAN = neurocan; PDGFA = platelet-derived growth factor alpha polypeptide; PNPLA3 = patatin-like phospholipase domain containing 3; SEM = standard error of the mean; SNP = single nucleotide polymorphism.

between rs738409 and visceral adipose tissue volume on liver fat content (41). Each copy of the rs738409 G-allele and 100 cm^3/150 mm slice visceral adipose tissue decreased liver attenuation value by 2.68 ± 0.35 Hounsfield units ($P < 0.01$). Further studies are needed to validate the combined effect of obesity and the genetic variant.

Because many studies reported that increased ALT activities were the most common abnormality in patients with nonalcoholic steatohepatitis (a progressive stage of NAFLD) (42), we detected the relations between the 7 SNPs and ALT level. We found the significant association between the G-allele of rs738409 and ALT level, with the effect of 1.09 IU/L per G-allele (95% CI 0.25–1.93, $P = 0.011$). The result was consistent with 1 GWAS study that identified its association with elevated ALT level in Europeans and Indian Asians (43). The association was also replicated in independent candidate gene studies (24,44) and a meta-analysis across different populations (40). The study results indicated that the SNP is associated with the progressive stage of NAFLD in children.

Our study examined the cumulative effect of 7 SNPs. Subjects carrying 10 or more risk alleles had an OR of 4.76 (95% CI 1.22–18.59, $P = 0.025$) compared with subjects that carry 3 or fewer risk alleles. Further large-scale studies in different ethnic populations are needed to validate the association between these gene variants and NAFLD, and the functional studies should be performed to elucidate the mechanism of NAFLD in children, to prevent adult NAFLD and other related diseases.

The limitations of the present study should be noted. The first is the limited sample size and statistic power. Because no previous study reported the effects of these SNPs on NAFLD in Chinese children, we could only estimate the statistic power. Under an additive genetic model, at a significance level of $P < 0.05$, our study had >70% power to detect the assumed effect size (OR 1.5) for the 7 SNPs with effect allele frequencies ≥0.10. By comparing the effect allele frequencies in our study with those in previous GWAS studies (14–16), we found that 2 SNPs (*FDFT1* rs2645424 and *COL13A1* rs1227756) had moderate differentiation (F_{ST} between 0.05 and 0.15). We considered that the effect of these SNPs in Chinese may be lower than that in Europeans, which could not be determined with our sample size. Second, the case-control study design did not permit us to make conclusions about causality. Thirdly, the diagnosis of NAFLD was based on questionnaire and abdominal ultrasound examination, but not on histologic examination. Histologic examination is the standard for the diagnosis of NAFLD and its severity, but it is invasive. We could not perform histologic examination for children. The diagnostic criteria based on ultrasound examinations have, however, been previously proved to be able to differentiate mild, moderate, and severe steatosis (29). In our study, we used the levels of ALT, which increased with the severity of NAFLD and can be complementary to the criteria. Furthermore, the children were not tested for other forms of liver disease such as Wilson disease because this was a population study.

In conclusion, we found that the *PNPLA3* rs738409 G-allele was associated with NAFLD and had stronger association with moderate-to-severe steatosis in Chinese children. The association between the *PNPLA3* rs738409 and increased ALT levels was also validated. For the 7 SNPs identified by GWAS, children carrying 10

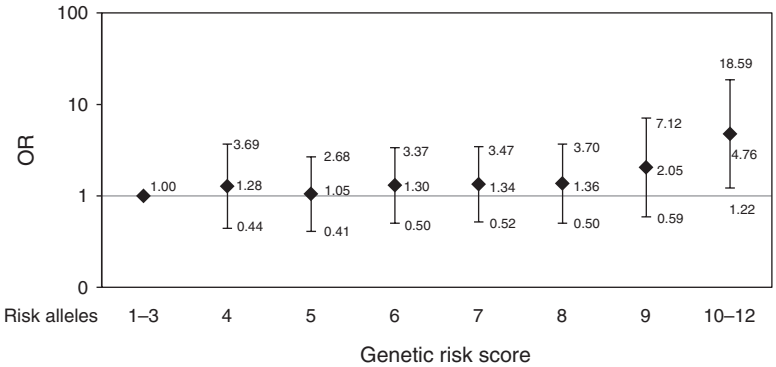

FIGURE 1. ORs and 95% CIs for the risk of NAFLD among children with different genetic risk scores. The score ≤3 was used as the reference in logistic regression analysis. On average, each additional effect allele was associated with increased risk of NAFLD (OR 1.13, 95% CI 1.002–1.28, $P = 0.046$). CI = confidence interval; NAFLD = nonalcoholic fatty liver disease; OR = odds ratio.

or more risk alleles were susceptible for NAFLD. The results provided evidence for identifying genetic factors of NAFLD and developing risk assessment and personalized medicine of NAFLD.

REFERENCES

1. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313–21.
2. Milic S, Stimac D. Nonalcoholic fatty liver disease/steatohepatitis: epidemiology, pathogenesis, clinical presentation and treatment. *Dig Dis* 2012;30:158–62.
3. Park SH, Jeon WK, Kim SH, et al. Prevalence and risk factors of non-alcoholic fatty liver disease among Korean adults. *J Gastroenterol Hepatol* 2006;21:138–43.
4. Fan JG, Farrell GC. Epidemiology of non-alcoholic fatty liver disease in China. *J Hepatol* 2009;50:204–10.
5. Chitturi S, Farrell GC, George J. Non-alcoholic steatohepatitis in the Asia-Pacific region: future shock? *J Gastroenterol Hepatol* 2004;19: 368–74.
6. Tominaga K, Kurata JH, Chen YK, et al. Prevalence of fatty liver in Japanese children and relationship to obesity. An epidemiological ultrasonographic survey. *Dig Dis Sci* 1995;40:2002–9.
7. Janssen I, Katzmarzyk PT, Boyce WF, et al. Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns. *Obes Rev* 2005;6:123–32.
8. Angulo P. Nonalcoholic fatty liver disease. *N Engl J Med* 2002;346: 1221–31.
9. Abdelmalek MF, Liu C, Shuster J, et al. Familial aggregation of insulin resistance in first-degree relatives of patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* 2006;4:1162–9.
10. Willner IR, Waters B, Patil SR, et al. Ninety patients with nonalcoholic steatohepatitis: insulin resistance, familial tendency, and severity of disease. *Am J Gastroenterol* 2001;96:2957–61.
11. Brouwers MC, van Greevenbroek MM, Cantor RM. Heritability of nonalcoholic fatty liver disease. *Gastroenterology* 2009;137:1536.
12. Schwimmer JB, Celedon MA, Lavine JE, et al. Heritability of non-alcoholic fatty liver disease. *Gastroenterology* 2009;136:1585–92.
13. Daly AK, Ballestri S, Carulli L, et al. Genetic determinants of susceptibility and severity in nonalcoholic fatty liver disease. *Expert Rev Gastroenterol Hepatol* 2011;5:253–63.
14. 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.
15. Chalasani N, Guo X, Loomba R, et al. Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. *Gastroenterology* 2010;139:1567–76.
16. Speliotes EK, Yerges-Armstrong LM, Wu J, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet* 2011;7:e1001324.
17. Yoneda M, Endo H, Mawatari H, et al. Gene expression profiling of non-alcoholic steatohepatitis using gene set enrichment analysis. *Hepatol Res* 2008;38:1204–12.
18. Miquilena-Colina ME, Lima-Cabello E, Sanchez-Campos S, et al. Hepatic fatty acid translocase CD36 upregulation is associated with insulin resistance, hyperinsulinaemia and increased steatosis in non-alcoholic steatohepatitis and chronic hepatitis C. *Gut* 2011;60:1394–402.
19. Hernaez R. Genetic factors associated with the presence and progression of nonalcoholic fatty liver disease: a narrative review. *Gastroenterol Hepatol* 2012;35:32–41.
20. Yang Z, Wen J, Tao X, et al. Genetic variation in the GCKR gene is associated with non-alcoholic fatty liver disease in Chinese people. *Mol Biol Rep* 2011;38:1145–50.
21. Peng XE, Wu YL, Lin SW, et al. Genetic variants in PNPLA3 and risk of non-alcoholic fatty liver disease in a Han Chinese population. *PLoS One* 2012;7:e50256.
22. Li X, Zhao Q, Wu K, et al. I148M variant of PNPLA3 confer increased risk for nonalcoholic fatty liver disease not only in European population, but also in Chinese population. *Hepatology* 2011;54:2275.
23. Li Y, Xing C, Tian Z, et al. Genetic variant I148M in PNPLA3 is associated with the ultrasonography-determined steatosis degree in a Chinese population. *BMC Med Genet* 2012;13:113.
24. Wang CW, Lin HY, Shin SJ, et al. The PNPLA3 I148M polymorphism is associated with insulin resistance and nonalcoholic fatty liver disease in a normoglycaemic population. *Liver Int* 2011;31:1326–31.
25. Lin YC, Chang PF, Chang MH, et al. Genetic variants in GCKR and PNPLA3 confer susceptibility to nonalcoholic fatty liver disease in obese individuals. *Am J Clin Nutr* 2014;99:869–74.
26. Lin YC, Chang PF, Hu FC, et al. A common variant in the PNPLA3 gene is a risk factor for non-alcoholic fatty liver disease in obese Taiwanese children. *J Pediatr* 2011;158:740–4.
27. Wang HJ, Zhang H, Zhang SW, et al. Association of the common genetic variant upstream of INSIG2 gene with obesity related phenotypes in Chinese children and adolescents. *Biomed Environ Sci* 2008;21:528–36.
28. Zeng MD, Fan JG, Lu LG, et al. Guidelines for the diagnosis and treatment of nonalcoholic fatty liver diseases. *J Dig Dis* 2008;9:108–12.
29. Graif M, Yanuka M, Baraz M, et al. Quantitative estimation of attenuation in ultrasound video images: correlation with histology in diffuse liver disease. *Invest Radiol* 2000;35:319–24.
30. Ji CY. Report on childhood obesity in China (1)—body mass index reference for screening overweight and obesity in Chinese school-age children. *Biomed Environ Sci* 2005;18:390–400.
31. Jie M, Tian-you W, Ling-hui M, et al. Development of blood pressure reference standards for Chinese children and adolescents. *Chin J Evid Based Pediatr* 2010;5:4–14.
32. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.
33. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution* 1984:1358–70.
34. Duan S, Zhang W, Cox NJ, et al. FstSNP-HapMap3: a database of SNPs with high population differentiation for HapMap3. *Bioinformation* 2008;3:139–41.
35. Balloux F, Lugon-Moulin N. The estimation of population differentiation with microsatellite markers. *Mol Ecol* 2002;11:155–65.
36. Gorden A, Yang R, Yerges-Armstrong LM, et al. Genetic variation at NCAN locus is associated with inflammation and fibrosis in non-alcoholic fatty liver disease in morbid obesity. *Hum Hered* 2013; 75:34–43.
37. He S, McPhaul C, Li JZ, et al. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem* 2010;285:6706–15.
38. Kumari M, Schoiswohl G, Chitraju C, et al. Adiponutrin functions as a nutritionally regulated lysophosphatidic acid acyltransferase. *Cell Metab* 2012;15:691–702.
39. Li JZ, Huang Y, Karaman R, et al. Chronic overexpression of PNPLA3I148M in mouse liver causes hepatic steatosis. *J Clin Invest* 2012;122:4130–44.
40. 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.
41. Graff M, North KE, Franceschini N, et al. PNPLA3 gene-by-visceral adipose tissue volume interaction and the pathogenesis of fatty liver disease: the NHLBI family heart study. *Int J Obes (Lond)* 2013;37:432–8.
42. Diehl AM, Goodman Z, Ishak KG. Alcohollike liver disease in non-alcoholics. A clinical and histologic comparison with alcohol-induced liver injury. *Gastroenterology* 1988;95:1056–62.
43. 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.
44. Sookoian S, Castano GO, Burgueno 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.