HEPATOLOGY
Official Journal of the American Association for the Study of Liver Diseases

Genetic Variation in Transmembrane 6 Superfamily Member 2 and the Risk of Nonalcoholic Fatty Liver Disease and Histological Disease Severity

Silvia Sookoian,[1]* Gustavo O. Castaño,[2] Romina Scian,[1,4] Pablo Mallardi,[3] Tomas Fernández Gianotti,[4] Adriana L. Burgueño,[4] Julio San Martino,[3] and Carlos J. Pirola[4]*

We explored the role of transmembrane 6 superfamily member 2 (*TM6SF2*) rs58542926 C/T nonsynonymous (p.Glu167Lys) variant in genetic susceptibility to nonalcoholic fatty liver disease (NAFLD) and disease severity. A total of 361 individuals (135 control subjects and 226 patients with histologically proven NAFLD) were included in a sample with 97% power for the additive genetic model. A discrete trait analysis of NAFLD showed that rs58542926 was associated with a modest risk of fatty liver ($P = 0.038$; odds ratio [OR]: 1.37; 95% confidence interval [CI]: 1.02-1.84); nevertheless, conditioning on patatin-like phospholipase domain-containing 3 (*PNPLA3*)-rs738409 abolished this effect. We did not observe an interaction between rs738409 and rs58542926 variants on the risk of NAFLD. We observed a significant association of rs58542926 and disease severity ($P = 0.027$), but not lobular inflammation or fibrosis; rs58542926 was not associated with levels of liver enzymes. An allelic test showed that the T (Lys167) allele was significantly associated with disease progression ($P = 0.021$; OR, 1.66; 95% CI: 1.08-2.55). A significant association was found with the histological degree of liver steatosis (β, 0.15; standard error: 0.06; $P = 0.0299$) that was independent of rs738409. Homozygous carriers of the C (Glu167) allele showed increased risk for cardiovascular disease. TM6SF2 protein expression was decreased markedly in liver of NAFLD patients, compared to controls. In addition, TM6SF2 immunoreactivity was reduced in subjects carrying at least one copy of the T allele, consistent with a difference in liver allele-specific transcript abundance. *Conclusion*: rs58542926 is a low-frequency variant with a modest effect on NAFLD, suggesting that carriers of the T allele are slightly more likely to accumulate fat in the liver and develop nonalcoholic steatohepatitis than those without. TM6SF2 appears to play a significant role in disease biology. (HEPATOLOGY 2015;61:515-525)

Abbreviations: Ab, antibody; ALT, alanine aminotransferase; ANCOVA, analysis of covariance; AST, aspartate aminotransferase, BMI, body mass index; cDNA, complimentary DNA; CI, confidence interval; CRP, C-reactive protein; CVD, cardiovascular disease; ER, endoplasmic reticulum; gDNA, genomic DNA; GWAS, genome-wide association studies; ^{1}H-MRS, proton magnetic resonance spectroscopy; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostasis model assessment for IR; HTGC, hepatic triglyceride content; IHC, immunohistochemistry; IR, insulin resistance; LDL-C, low-density lipoprotein cholesterol; MAF, minor allele frequency; MetS, metabolic syndrome; mRNA, messenger RNA; NAFL, nonalcoholic fatty liver; NAFLD, nonalcoholic fatty liver disease; NAS, NAFLD Activity Score; NASH, nonalcoholic steatohepatitis; OR, odds ratio; PNPLA3, patatin-like phospholipase domain-containing 3; SD, standard deviation; SE, standard error; SNP, single-nucleotide polymorphism; TC, total cholesterol; TE, transient elastography; TG, triglyceride; TM6SF2, transmembrane 6 superfamily member 2; US, ultrasound..

From the [1]Department of Clinical and Molecular Hepatology, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires–National Scientific and Technical Research Council (CONICET), Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina; [2]Liver Unit, Medicine and Surgery Department, Hospital Abel Zubizarreta, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina; [3]Pathology Department, Hospital Diego Thompson, San Martin, Buenos Aires, Argentina; and [4]Department of Molecular Genetics and Biology of Complex Diseases, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires–National Scientific and Technical Research Council (CONICET), Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

Received July 20, 2014; accepted October 5, 2014.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.27556/suppinfo.

This study was partially supported by grants PICT 2010-0441 and PICT 2012-0159 (Agencia Nacional de Promoción Científica y Tecnológica) and UBACYT CM04 (Universidad de Buenos Aires).

**Drs. Sookoian and Pirola share joint senior authorship.*

Nonalcoholic fatty liver (NAFL) disease (NAFLD) is a common chronic liver disease whose prevalence has reached global epidemic proportions.[1] As with other complex disorders, NAFLD develops from a combination of many risk factors, including genetic susceptibility, given by the effect of multiple genes and environmental insults.[2] Although this concept indicates that the genetic basis of NAFLD is polygenic and multifactorial, there is compelling and unequivocal evidence that the rs738409 variant (encoding an amino acid substitution p.Ile148Met) of *PNPLA3* (patatin-like phospholipase domain-containing 3; also known as adiponutrin or calcium-independent phospholipase A2-epsilon) is a strong modifier of the natural history of NAFLD, by modulating liver fat deposition[3] and disease severity and progression, and accounting for approximately 5% of disease variance.[4]

Recently, scientists from the Dallas Heart Study employed an exome-wide approach to elucidate further variants involved in liver fat accumulation; this technology included the exploration of a large number of functional exonic variants (247,870 markers) in a chip.[5] Interestingly, a nonsynonymous variant located in the *TM6SF2* (transmembrane 6 superfamily member 2) gene, rs58542926 encoding an amino acid substitution p.Glu167Lys (E167K), was significantly associated with hepatic triglyceride content (HTGC), as measured by proton magnetic resonance spectroscopy (^{1}H-MRS). The investigators showed that the effect of rs58542926 on HTGC was independent of the effect mediated by rs738409, obesity, or insulin resistance (IR), as assessed by the homeostasis model assessment index for IR (HOMA-IR), or alcohol intake.[5] In addition, Kozlitina et al. reported that the *TM6SF2* variant was associated with a significant increase in serum alanine aminotransferase (ALT) activity, but not in aspartate aminotransferase (AST). The investigators replicated this finding in two large, population-based studies: the Dallas Biobank and the Copenhagen City Heart Study. Thus, the investigators inferred that rs58542926 is associated with increased hepatic injury.[5]

A recent hospital-based study on patients with NAFLD proven by liver biopsy in the absence of controls with appropriate phenotype characterization compared the genotypes counts of rs58542926 in patients versus those in the 1,000 Genomes from the European Caucasian population and concluded that the variant is associated with NAFLD, but not with histological steatosis, suggesting that, if any, the effect of rs58542926 on liver fat accumulation is of relatively small size.[6] In addition, Liu et al. found that rs58542926 was associated with necroinflammation score, but, unfortunately, they were unable to replicate this finding in their larger validation cohort.[6] On the contrary, Liu et al. observed that rs58542926 was significantly associated with histological fibrosis in the two explored cohorts of patients by adopting an additive model.[6]

However, Wong et al. reported that the *TM6SF2* variant was not associated either with liver fat accumulation assessed by ^{1}H-MRS or with liver fibrosis assessed by transient elastography (TE) in a large, community-based study from China.[7]

Hence, whereas the role of *PNPLA3* on the pathogenesis and genetic risk of NAFLD appears to be straightforward, the involvement of the newly NAFLD-associated gene variant, *TM6SF2*-rs58542926, on the susceptibility of the disease is controversial. In this study, we performed a hospital-based adult case-control association study to explore the association between rs58542926 and NAFLD susceptibility and evaluate further the association between the variant and histological disease severity in patients with NAFLD proven by liver biopsy. A functional study to understand the role of the variant on gene and protein expression was also included.

Patients and Methods

Patients and Control Subjects: Selection Criteria. The study included 361 unrelated individuals, of which 135 were control subjects and 226 were patients who have histopathologically proven features of NAFLD (including 96 with simple steatosis [NAFL], and 130 with nonalcoholic steatohepatitis [NASH]).

Address reprint requests to: Silvia Sookoian, M.D., Ph.D., Instituto de Investigaciones Medicas, IDIM-CONICET, Combatientes de Malvinas 3150, CABA-1427, Argentina. E-mail: sookoian.silvia@lanari.fmed.uba.ar; fax: 54-11-4523 8947 or Carlos J. Pirola, Ph.D., Instituto de Investigaciones Medicas, IDIM-CONICET Combatientes de Malvinas 3150, CABA-1427, Argentina. E-mail: pirola.carlos@lanari.fmed.uba.ar; fax: 54-11-4523 8947.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.27556

Potential conflict of interest: Nothing to report.

The study population included participants involved in the first report, demonstrating that rs738409 modulates disease severity.[8]

Secondary causes of steatosis, including alcohol abuse (≥30 g of alcohol daily for men and ≥20 g for women), total parenteral nutrition, hepatitis B and hepatitis C virus infection and the use of drugs known to precipitate steatosis were excluded. In addition, patients with any of the following diseases were excluded from participation: autoimmune liver disease; metabolic liver disease; Wilson's disease; and α-1-antitrypsin deficiency.

Control subjects were selected from subjects attending our hospital for check-up purposes whose sex matched the NAFLD patients. In addition to the standard heath examination, all control individuals were subjected to a liver ultrasound (US) examination. They were included in the study if they did not have evidence of fatty change or biochemical abnormalities. Furthermore, control subjects were confirmed not to have any of the features of metabolic syndrome (MetS), as defined by the National Cholesterol Education Program Adult Treatment Panel III, and did not abuse alcohol.

Case participants and controls were selected during the same study period from the same population of patients attending our institution, and all share the same demographic characteristics (occupation, educational level, place of residence, and ethnicity).

All investigations performed were conducted in accord with the guidelines of the 1975 Declaration of Helsinki. Written consent from individuals was obtained in accord with the procedures approved by the ethical committee of our institution.

Physical, Anthropometric, and Biochemical Evaluation. Health examinations included anthropometric measurements, a questionnaire on health-related behaviors, biochemical determinations, and assessment of risk of cardiovascular disease (CVD). Details are provided in the Supporting Information.

CVD Assessment and Characterization of a Proatherogenic Phenotype. CVD risk was evaluated by exploring the total cholesterol (TC)/high-density lipoprotein cholesterol (HDL-C) ratio[9] and specific biomarkers of atherogenesis and systemic inflammation; complete details are provided in the Supporting Information.

Liver Biopsy and Histopathological Evaluation. Liver biopsy was performed before any intervention with US guidance and a modified 1.4-mm-diameter Menghini needle (Hepafix, Braun, Germany) under local anesthesia on an outpatient basis. A portion of each liver biopsy specimen was routinely fixed in 40 g/L of formaldehyde (pH 7.4), embedded in paraffin, and stained with hematoxylin and eosin, Masson trichrome, and silver impregnation for reticular fibers. All biopsies were at least 3 cm in length and contained a minimum of eight portal tracts. Degree of steatosis was assessed according to the system developed by Kleiner et al., based on the percentage of hepatocytes containing macrovesicular fat droplets.[10] NASH[11] and NAFLD Activity Score (NAS)[10] were defined as reported previously; a NAS threshold of 5 was used for further comparisons with variables of interest, and NASH was defined as steatosis plus mixed inflammatory-cell infiltration, hepatocyte ballooning and necrosis, glycogen nuclei, Mallory's hyaline, and any stage of fibrosis, including absent fibrosis.[10]

Expression of Liver TM6SF2 Assessed by Immunohistochemistry. Immunostaining for TM6SF2 was performed on liver specimens of NAFLD patients and the controls previously included in paraffin. Control liver specimens were obtained by percutaneous liver biopsy from sex- and age-matched adults with mildly elevated serum liver enzyme activity for whom all causes of liver disease were ruled out.

Slides were incubated with a dilution of 1:100 of rabbit polyclonal antibody (Ab) specific for Human anti-TM6SF2 Ab (TM6SF2 Antibody—C-terminal region—ARP44400_P050; Aviva Systems Biology, San Diego, CA).

The pathologists (P.M. and J.S.M.) evaluated the TM6SF2 immunostaining semiquantitatively in a blinded fashion regarding any of the histological and clinical characteristics of the patients; the degree of interobserver variability was determined by K statistics. To confirm the specificity of the binding affinity of the TM6SF2 Ab, we performed a blocking peptide competition protocol. Complete details are provided in the Supporting Information.

Exploration of Allele-Specific Transcript Liver Expression Using Single-Nucleotide Polymorphism Genotyping of Complimentary DNA. To explore whether rs58542926-allele–specific expression contributes to TM6SF2 liver expression, we quantified both liver complimentary DNA (cDNA) and genomic DNA (gDNA) from 10 patients heterozygous for the variant. Complete details are provided in the Supporting Information.

Genotype and Association Analysis, and Power and Sample-Size Calculation. Genetic analyses were done on gDNA extracted from white blood cells. Genotyping of the TM6SF2 rs58542926 was performed using a TaqMan genotyping assay (dbSNP rs58542926

Table 1. Clinical and Biochemical Characteristics of Control Subjects and Patients With NAFLD

Variables	Control Subjects	NAFL	NASH	*P* Value*	*P* Value†	*P* Value‡
Demographic and lifestyle factors						
No. subjects	135	96	130	-	-	-
Female/male, n	84/51	53/43	88/42	NS	NS	NS
Age, years	48 ± 12	53 ± 13	51 ± 12	0.01	NS	0.04
Smoking habit, cigarettes/day	1.8 ± 4.3	5.9 ± 12.0	2.9 ± 7.7	NS	NS	NS
Physical activity, h/week	1.5 ± 2.7	1.6 ± 4.9	1.7 ± 6.0	NS	NS	NS
Metabolic risk factors						
BMI, kg/m^2	25 ± 4	31.5 ± 6.0	33.4 ± 6.0	1.0×10^{-8}	0.008	1.0×10^{-8}
Waist circumference, cm	84 ± 15	102 ± 15	107.0 ± 12.5	1.0×10^{-8}	0.009	1.0×10^{-8}
Waist/hip ratio	0.84 ± 0.09	0.9 ± 0.07	0.9 ± 0.08	1.0×10^{-8}	NS	1.0×10^{-8}
Fasting plasma glucose, mg/dL	83.4 ± 13.0	97.7 ± 19.0	130 ± 128	1.0×10^{-8}	1.0×10^{-8}	1.0×10^{-8}
Fasting plasma insulin, μU/ml	6.7 ± 4.7	12.7 ± 9.0	16.0 ± 10.4	1.0×10^{-8}	0.002	1.0×10^{-8}
HOMA-IR	1.4 ± 1.0	3 ± 2	5.0 ± 6.7	1.0×10^{-8}	0.00004	1.0×10^{-8}
CVD risk						
SABP, mmHg	115 ± 14	125 ± 14	127.4 ± 15.0	0.00002	NS	1.0×10^{-8}
DABP, mmHg	72.0 ± 9.5	76.8 ± 11.0	78.7 ± 10.7	0.002	NS	1.0×10^{-8}
Leukocyte count, cells/mm^3	6,702 ± 2,181	7,486 ± 2,025	7,826 ± 2,254	0.07	NS	0.004
CRP	6.3 ± 4.3	6.1 ± 3.3	7.5 ± 5.0	NS	NS	NS
sICAM-1	360 ± 184	515 ± 231	660 ± 344	0.001	0.05	1.0×10^{-8}
PAI-1	18,873 ± 21,890	23,221 ± 18,979	24,500 ± 16,620	NS	NS	0.00005
Soluble CD40 ligand	809 ± 398	1,671 ± 6,188	1,093 ± 4,923	NS	NS	0.05
Resistin	7,221 ± 9,847	8,216 ± 9,711	8,423 ± 5,189	NS	NS	NS
TC, mg/dL	206 ± 40	207 ± 50	212 ± 43	NS	NS	NS
HDL-C, mg/dL	56 ± 16	53 ± 25	49.5 ± 13.0	0.02	NS	0.0001
LDL-C, mg/dL	124 ± 38	125 ± 48	126 ± 43	NS	NS	NS
TGs, mg/dL	122 ± 79	146 ± 73	200 ± 124	0.007	0.001	1.0×10^{-8}
Total cholesterol/HDL-C ratio	3.7 ± 3.0	3.6 ± 7.7	4.0 ± 1.75	NS	NS	0.003
Uric acid, mg/dL	3.3 ± 1.4	4.6 ± 2.4	4.9 ± 2.0	0.00001	NS	1.0×10^{-8}
Liver-related phenotype						
ALT, U/L	17 ± 9	57.8 ± 64.8	70.8 ± 56.0	1.0×10^{-8}	0.003	1.0×10^{-8}
AST, U/L	19 ± 9	36.0 ± 17.5	50 ± 34	1.0×10^{-8}	0.00004	1.0×10^{-8}
GGT, U/L	24.3 ± 22.0	66 ± 58	87.5 ± 89.0	0.00002	0.06	1.0×10^{-8}
ALP, U/L	139 ± 58	230 ± 110	234 ± 113	0.00001	NS	0.00002
Histological features						
Degree of steatosis, %	-	47 ± 25	59 ± 21	-	0.0004	-
Lobular inflammation (0-3)	-	0.7 ± 0.6	1.3 ± 0.6	-	1.0×10^{-8}	-
Portal inflammation (0-2)	-	0.0 ± 0.1	1.5 ± 0.7	-	1.0×10^{-8}	-
Hepatocellular ballooning (0-2)	-	0.0 ± 0.2	0.8 ± 0.5	-	1.0×10^{-8}	-
Fibrosis stage	-	0.0 ± 0.1	1.4 ± 1.3	-	1.0×10^{-8}	-
NAS	-	2.7 ± 1.1	5.8 ± 1.45		1.0×10^{-8}	

Results are expressed as mean ± SD. *P* value stands for statistical significance using Mann-Whitney's *U* test, except for female/male proportion, where the *P* value stands for statistical significance using the chi-squared test.

*Comparisons between NAFL versus control subjects.

†Comparisons between NASH versus NAFL.

‡Comparisons between NASH versus controls.

Abbreviations: SABP and DABP, systolic and diastolic arterial blood pressure, respectively; sICAM-1, soluble intercellular adhesion molecule 1; PAI-1, plasminogen activator inhibitor 1; GGT, gamma-glutamyl transferase; ALP, alkaline phosphatase; NS, nonsignificant.

assay C__89463510_10, #4351379; Applied Biosystems, Foster City, CA), according to manufacturer's instructions. Genotyping of PNPLA3 rs738409 was performed using a high-throughput genotyping method, as previously described.[8] To ensure genotyping quality, we included DNA samples as internal controls, hidden samples of known genotype, and negative controls (water). The overall genotype completion rate was 100%.

To account for possible population stratification, we used a collection of 13 single-nucleotide polymorphisms (SNPs; rs6830727, rs12639788, rs1282807, rs1947745, rs7162312, rs12951674, rs7212346, rs1934869, rs9542666, rs11843545, rs9725124, rs2798659, and rs2199940) at different loci (located in chromosomes 4, 15, 17, 13, 1, and 3) and then analyzed the data with the Structure program (Version 2).[12]

We found no evidence of stratification in our sample because the cases and controls showed similar Q values and the Structure program assigned a similar distance to clusters with no further improvement in

Table 2. Multiple Regression Analysis of Serum Liver Enzymes (ALT and AST) as Dependent Variables and *TM6SF2*-rs58542926 Genotypes, Age, BMI, HOMA, and *PNPLA3*-rs738409 as Independent Variables

	ALT			AST		
Variables	$\beta \pm$ SE	B $\pm$ SE	*P* Level	$\beta \pm$ SE	B $\pm$ SE	*P* Level
rs58542926	0.12 $\pm$ 0.06	0.11 $\pm$ 0.06	0.06	0.058 $\pm$ 0.06	0.04 $\pm$ 0.04	0.37
Age	0.10 $\pm$ 0.06	0.006 $\pm$ 0.004	0.12	0.11 $\pm$ 0.06	0.005 $\pm$ 0.003	0.08
Log BMI	0.37 $\pm$ 0.06	1.45 $\pm$ 0.25	0.000001	0.34 $\pm$ 0.06	1.03 $\pm$ 0.20	0.000001
Log HOMA	0.10 $\pm$ 0.08	0.09 $\pm$ 0.07	0.21	0.04 $\pm$ 0.08	0.03 $\pm$ 0.06	0.57
rs738409	0.17 $\pm$ 0.06	0.10 $\pm$ 0.04	0.005	0.17 $\pm$ 0.06	0.07 $\pm$ 0.03	0.008

Values are beta $\pm$ SE. A logarithm (log) transformation was applied to traits, including ALT and AST values with non-normal distributions.

the fitting model by adding up to four clusters (the ln of likelihood was maximum for K = 1). Moreover, all the participants in this study self-reported a Caucasian ethnicity as a surrogate of ancestry, which is consistent with the minor allele frequency (MAF) found.

Based on previous reports of the association between rs58542926 and NAFLD,[5,6] we used the additive genetic model of inheritance, unless otherwise indicated. To explore the risk of CVD, we used the dominant model.[13] Using the CaTS power calculator for genetic association studies[14] and assuming a prevalence of NAFLD of 0.30 and a rs58542926 MAF of 0.08, our sample had 97% power for the additive genetic model and 94% power for the dominant one.

Statistical Analysis. Quantitative data are expressed as mean $\pm$ standard deviation (SD), unless otherwise indicated. Given that a significant difference was observed between groups in most of the variables and the distribution was significantly skewed in most cases, we chose to be conservative and assessed the differences between groups using nonparametric Mann-Whitney's *U* or Kruskal-Wallis' tests. Cochran-Armitage's test for trend was used in the categorical data analysis to assess the presence of association between the variant and disease severity and a regression analysis for an ordinal multinomial distribution (Probit as the Link function), with disease severity as the dependent (response) variable coding controls; NAFL and NASH subjects as 0, 1, and 2, respectively; age, HOMA, and body mass index (BMI) as continuous predictor variables; and sex and rs738409 and rs58542926 genotypes (0,1,2) as grouping variables. Moreover, logistic regression analysis was included for the evaluation of the association between genotypes and histological disease severity (NAS, ballooning, fibrosis, and inflammation: present coded as 1 or absent coded as 0). To assess the association between genotypes with NAFLD or quantitative traits, such as ALT and AST, we used a chi-square test and logistic regression or analysis of covariance (ANCOVA) and multiple regression, adjusting for covariables, such as age, HOMA, BMI, and rs738409. For ordinal multinomial analysis, logistic analysis, or ANCOVA, we adjusted for covariables that were not normally distributed through log transformation. Correlation between two variables was done using Spearman's rank-correlation test. The CSS/Statistica program package (version 6.0; StatSoft, Tulsa, OK) was used in these analyses.

Results

Table 1 shows the details about the physical, anthropometric, biochemical, and CVD evaluation of patients and controls. NAFLD patients showed most of the risk factors of MetS, including elevated BMI, waist-hip ratio, fasting glucose, insulin and HOMA-index, and CVD risk.

Association Between the TM6SF2 Variant and NAFLD: The rs58542926 Is Associated With a Modest Risk of Fatty Liver. In controls, frequencies of the C (Glu167) and the T (Lys167) alleles were 94.5% and 5.5%, respectively, and distribution of genotypes was in Hardy-Weinberg's equilibrium ($P = 0.5$). C/T alleles in the forward strand correspond to G/A alleles in the reverse coding strand; the C (or G) allele is the ancestral one.

In the whole population, frequencies of the C allele and T allele were 91.4% and 8.6%, respectively; distribution of genotypes was also in Hardy-Weinberg's equilibrium ($P = 0.11$).

Discrete trait analysis of NAFLD showed that rs58542926 was associated with a modest risk of fatty liver in the additive model ($P = 0.038$; odds ratio [OR]: 1.37; 95% confidence interval (CI): 1.02-1.84); however, after incorporation of *PNPLA3*-rs738409, age, HOMA-IR, and BMI into multiple logistic regression analysis, the *P* value was not significant. Furthermore, we did not observe any interaction between effects of the *PNPLA3*−738409 and *TM6SF2*-rs58542926 variants on NAFLD. Finally, univariate or multiple regression analysis showed that

Table 3. Genotypic Test for Trend: *TM6SF2*-rs58542926 Genotypes According to Disease Severity

Disease Status	EE	EK	KK	*P* Value
Control subjects	120	15	0	0.027
NAFL	79	16	1	
NASH	105	21	4	

P value stands for statistical significance using Mantel-Haenszel's chi-square (degree of freedom = 1).
EE, homozygous for the allele encoding Glu167 (genotype CC); EK, heterozygous (genotype CT); KK, homozygous for the allele encoding Lys167 (genotype TT).

TM6SF2-rs58542926 was not associated with either serum ALT or AST values (Table 2).

Association Between the TM6SF2 Variant and Disease Severity: The rs58542926 Seems to Have a Small Effect on the Disease Progression Without Effect on Liver Fibrosis or Necroinflammation. In the analysis of genotypic test for trend, including control subjects and patients with NAFL and NASH, we observed a significant association of rs58542926 and the disease severity (Mantel-Haenszel's chi-squared: 4.89; $P = 0.027$; Table 3). An allelic association test showed that the T (Lys167) allele was significantly associated with disease progression (Cochran-Armitage's test for trend, chi-squared: 5.3; $P = 0.021$; generalized OR: 1.66; 95% CI: 1.08-2.55).

Nevertheless, analysis of rs58542926 and NAFLD histological stages (NASH vs. NAFL as the reference group) did not show a significant association (Cochran-Armitage's test for trend, chi-squared: 0.33; $P = 0.563$; generalized OR: 1.13; 95% CI: 0.58-2.20). Likewise, we observed a lack of association between the variant and NAS score (NAS score ≤ 5 vs. NAS score >5: OR, 1.16; 95% CI: 0.53-2.53; $P = 0.72$).

In addition, by grouping the dependant variable according to severity of liver fibrosis as absent or mild (F0-F1) and moderate or severe (F2-F3), we did not observe a significant association with rs58542926 (OR, 0.95; 95% CI: 0.66-1.36; $P = 0.77$). Similarly, neither lobular inflammation (OR, 0.85; 95% CI: 0.60-1.20; $P = 0.34$) or hepatocellular ballooning (OR, 1.08; 95% CI: 0.78-1.50; $P = 0.64$) were associated with the variant. Similar results were observed in the allelic test for trend that is independent of the genetic model (data not shown).

Notably, the degree of liver steatosis, as evaluated by liver biopsy and expressed as percentage of hepatocytes infiltrated by fat, was significantly associated with rs58542926 (β, 0.15; standard error [SE]: 0.06; $P = 0.0299$), and this association was independent of *PNPLA3*-rs738409 genotypes, age, sex, and BMI.

Association Between the TM6SF2 Variant and CVD: The Paradox of Carrying the NAFLD (T, Lys167) or the CVD (C, Glu167) Risk Allele. A recent genome-wide coding variation study that assessed variants for association with circulating lipid levels found that rs58542926 was significantly associated with plasma total and low-density lipoprotein cholesterol (LDL-C), and triglyceride (TG).[13] Notably, carriers of the C allele (Glu167) had abnormal lipid levels and an increased risk of myocardial infarction.[13] Similar results were observed recently by Kozlitina et al.[5] Thus, in our population, we explored whether subjects homozygous for the allele encoding Glu167 (genotype CC) have more risk factors for CVD than carriers of the T allele (heterozygous CT + homozygous for the allele encoding Lys167, genotype TT); the low number of homozygous for the T allele and visual inspection of the data justified this model. Interestingly, patients with the CC genotype, in comparison with those with CT + TT genotypes, had higher levels of C-reactive protein (CRP), TC, and sCD40 and a higher index of cardiovascular risk, even though the TM6SF2 variant was not associated with IR or obesity (Table 4).

The Effect of rs58542926 on NAFLD Biology: TM6SF2 Expression Is Significantly Decreased in the Liver of Patients With NAFLD and the Variant Might Regulate Liver Transcript and Protein Expression in an Allele-Specific Manner. Kozlitina et al. showed in vitro that murine hepatoma cells expressing Lys167-TM6SF2 protein have reduced expression levels, compared to the wild type.[5] Nevertheless, the putative impact of the rs58542926 genotypes on liver expression of TM6SF2 protein in patients with NAFLD is unknown.

Using immunohistochemistry (IHC), we explored the level of TM6SF2 liver expression according to disease severity (control liver, NAFL, and NASH) and the variant genotypes, and we observed that, compared to normal liver, TM6SF2 immunoreactivity was significantly decreased in liver tissue of patients with NAFLD, regardless of disease severity (Fig. 1A). Moreover, degree of liver steatosis was significantly and negatively correlated with the amount of the TM6SF2 immunoreactivity product (Spearman's R: -0.81; $P < 0.01$); likewise, NAS score was significantly and negatively correlated with TM6SF2 immunoreactivity (R, -0.72; $P < 0.04$). Conversely, analysis of TM6SF2 immunostaining and fibrosis score did not show a significant association.

Notably, statistical analysis focused on the rs58542926 genotypes revealed significantly reduced TM6SF2 immunoreactivity in subjects carrying at least

Table 4. The *TM6SF2* Variant and the Risk of CVD

Traits Associated With CVD	EE	EK+KK	*P* Value
Age, years	48.5 ± 12.7	50.3 ± 12.1	NS
BMI, kg/m^2	29.8 ± 6.5	30.2 ± 6.1	NS
Waist circumference, cm	97.6 ± 17.9	97.3 ± 16.3	NS
Waist/hip ratio	0.91 ± 0.09	0.98 ± 0.08	NS
SABP, mmHg	120.9 ± 15.7	124.7 ± 14.9	NS
DABP, mmHg	74.8 ± 10.6	77.4 ± 10.8	NS
CRP	7.0 ± 4.5	5 ± 3	0.03
Leukocyte count, cells/mm^3	7,244 ± 2,226	7,519 ± 2,200	NS
sICAM-1	505 ± 290	510 ± 322	NS
PAI-1	21,941 ± 20,304	18,633 ± 17,012	NS
Soluble CD40 ligand	1,209 ± 517	322 ± 535	0.01
Resistin	7,701 ± 5,465	8,372 ± 5,390	NS
TC, mg/dL	212 ± 44	195 ± 47	0.02
HDL-C, mg/dL	52.1 ± 18.7	54 ± 13	NS
LDL-C, mg/dL	127 ± 43	116 ± 38	NS
TGs, mg/dL	165.7 ± 109	141.4 ± 67.0	NS
TC/HDL-C ratio	4.0 ± 1.76	3.0 ± 1.60	0.004
Uric acid, mg/dL	4.3 ± 2.1	4.4 ± 1.9	NS
Fasting plasma glucose, mg/dL	104.6 ± 86.1	100.3 ± 27.9	NS
Fasting plasma insulin, μU/mL	11.4 ± 9.5	11.7 ± 6.8	NS
HOMA-IR	3.11 ± 4.7	2.98 ± 2.0	NS

Results are expressed as mean ± SD. The TC/HDL-C ratio was used as a measure of CVD risk. The association analysis was done on the dominant model of inheritance.[13] *P* value stands for statistical significance using Mann-Whitney's *U* test.

Abbreviations: SABP and DABP, systolic and diastolic arterial blood pressure, respectively; sICAM-1, soluble intercellular adhesion molecule 1; PAI-1, plasminogen activator inhibitor 1; NS, nonsignificant.

EE, homozygous for the allele encoding Glu167 (genotype CC); EK, heterozygous (genotype CT); KK, homozygous for the allele encoding Lys167 (genotype TT).

one copy of the NAFLD-risk T allele ($P < 0.01$, Mann-Whitney's test; Fig. 1B,D,E). The dominant model for the risk allele, which assumes that having one or more copies of the T allele increases disease risk compared to C, was used because of the low frequency of the T allele.

Finally, we observed that the immunoreactivity product was confined to hepatocytes, showing a primary cytoplasmic staining pattern (Fig. 1C-F); immunoreactivity in zones 1 and 3 were highly and significantly correlated (R, 0.80; $P < 0.01$). The kappa test for interobserver variation showed excellent agreement (kappa, 0.75). Supporting Fig. 1 illustrates the results of the blocking peptide competition protocol showing that the signal was highly specific.

To assess the effect of the p.Glu167Lys substitution on liver gene expression of *TM6SF2*, we quantified the relative allele-specific messenger RNA (mRNA) abundance in liver of heterozygous subjects normalized by the corresponding gDNA. Remarkably, allelic differences in gene expression were observed, owing to the fact that the liver C-allele mRNA abundance was 1.79-fold (± 1.18 SD) greater than the T-allele mRNA ($P = 0.0277$, nonparametric Wilcoxon's paired test).

Discussion

In this study, we observed that rs58542926, located in the *TM6SF2* locus, was associated with a small effect on the odds of having NAFLD, which, after conditioning on the *PNPLA3-rs*738409 and metabolic risk factors, including obesity and IR, was abolished. In addition, the rs58542926 T allele (Lys167) was associated with disease severity, suggesting that the variant might have a small effect on the risk of NASH. In our sample, rs58542926 was not associated either with serum levels of transaminases or with individual components of the NAS score, including necroinflammation and ballooning, or with liver fibrosis, except for a significant association with degree of histological steatosis. Consistent with data previously reported in a large, genome-wide association study (GWAS),[13] we observed that the rs58542926-C (Glu167) allele was significantly associated with TC and sCD40 levels and with increased risk of CVD.

Overall, these findings deserve several specific comments. First, the divergent results among published studies, as summarized in Table 5, suggest that the magnitude of the effect of the *TM6SF2* locus on the development of NAFLD is inconclusive. Unlike the overwhelming and straightforward evidence of the role of *PNPLA3* in the modulation of NAFLD and disease progression in different populations around the world,[4] the initial finding of the effect of rs58542926 on the genetic risk of fatty liver was either not replicated in Asians or hardly replicated in Caucasians (Table 5). A note of caution should be added because the lack of well-characterized controls in published studies might explain some discordant results. For instance, Liu et al. compared genotypic frequencies of NAFLD patients with data inferred from the 1000 Genomes Project, which includes genotyping of subjects with no associated medical or phenotype data.[6] This strategy, which may be used in an exploratory phase, is not recommended in replication or confirmatory studies.[15] Furthermore, there are also controversial results on the effect on liver fibrosis (Table 5), given that the study that found a positive association with rs58542926 failed to show a strong effect on disease severity, including necroinflammation.[6] Liu et al. speculated that *TM6SF2* drives NAFLD-associated hepatic fibrosis without affecting TG accumulation,[6] a hypothesis that essentially conspires against the two-hit theory, an explanation given by the same investigators to illustrate that NAFLD develops from a "multiple-hit" process, which begins with liver fat accumulation and continues with subsequent insults to developing

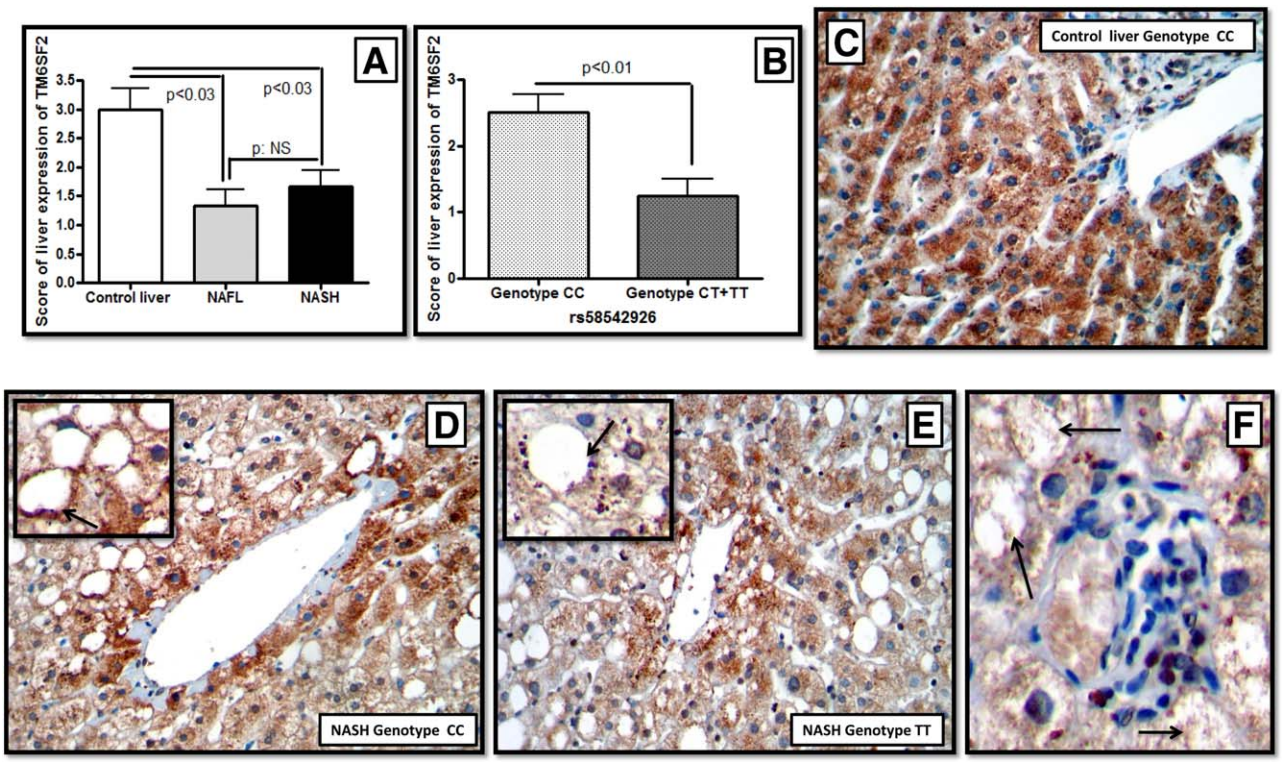

Fig. 1. Functional assessment of the impact of *TM6SF2* on NAFLD pathobiology using IHC. (A) Scores on liver protein expression of *TM6SF2* evaluated using IHC according to disease status (control liver, NAFL, and NASH; n = 4 per group). Statistical differences between groups were evaluated using the Mann-Whitney's *U* test. (B) Scores on liver protein expression of *TM6SF2* evaluated using IHC according to the grouping of rs58542926 genotypes by dominant model of inheritance. (C) Representative liver expression pattern of *TM6SF2* in control liver showing strong staining. (D and E) Representative liver expression pattern of *TM6SF2* in NASH patients carrying the rs58542926 CC or TT genotype, respectively; insert shows *TM6SF2* immunoreactivity surrounding the lipid droplets (black arrows). (F) Lobular inflammatory infiltrate showing negative immunoreactivity to *TM6SF2*; ballooned hepatocytes showing a decreased amount of the *TM6SF2* immunoreactivity product (black arrows). *TM6SF2* immunoreactivity was examined using light microscopy of liver sections; counterstaining was done with hematoxylin. *TM6SF2* immunostaining was significantly observed in the cytoplasm of hepatocytes showing a granular pattern. Original magnification: 400×. Abbreviation: NS, nonsignificant.

liver fibrosis.[16] While our study was under review, it was reported, by a large European study, that rs58542926 was associated with fibrosis severity, but the investigators observed that, after conditioning for NASH, this effect was abolished.[17]

Of note, we found that the variant was significantly associated with quantitative estimation of degree of hepatic steatosis. This finding might, in part, explain the association between rs58542926 and disease severity, given that it was shown that patients with severe steatosis are more likely to have steatohepatitis.[18] Similar findings were observed by Dongiovanni et al., who suggest that the effect of the variant on liver fibrosis is mediated by the increased lipid retention in hepatocytes, rather than by a direct "fibrogenic effect" of the risk allele.[17]

Second, unlike the common *PNPLA3*-rs738409 variant that has a MAF of ~0.30, *TM6SF2*-rs58542926 has a MAF of 0.06 in the Caucasian population. In our study, we did observe a whole-sample MAF of 0.086 and a MAF of 0.055 in the control group; similar findings have been reported thus far in other ethnicities (Table 5). Consequently, as shown in Table 5, the number of homozygous subjects for the T allele is very small, even in the larger studies; hence, none of the studies was able to demonstrate the putative effect of rs58542926 on a recessive model of inheritance, probably because none of them, including our own, had enough statistical power to do so. Therefore, the effect of rs58542926 on NAFLD deserves further investigation to confirm a putative recessive effect, given that the homozygous subjects for the T allele (Lys167) are rare in the whole population, with a frequency less than 1%. In this scenario, the variant is likely a minor contributor to the disease burden in the general population.

Third, *TM6SF2*-rs58542926 presents a clinical paradox; the C (Glu167) allele is associated consistently with increased CVD risk, and the T allele (Lys167) is associated putatively with NAFLD and NASH.

Table 5. A Summary of Studies That Explored the Association Between *TM6SF2*-rs58542926 and NAFLD and Related Traits

Main Findings	Kozlitina Et Al. 2014[5]	Liu Et Al. Discovery 2014[6]	Liu Et Al. Validation 2014[6]	Wong Et Al. 2014[7]	Dongiovanni Et Al. 2014[17]	Sookoian Et Al. (the Current Study)
Sample size	2,736	349	725	922	1,201	361
Study design	Population based	Patients with NAFLD*	Patients with NAFLD	Population based	Patients with NAFLD†	Case-control study
Ethnicity	Multiancestry	European (UK)	European (UK and Central Europe)	Chinese	European (Italy and Finland)	Argentinean (self-reported Central European ancestry)
Setting	General population (adult)	Hospital based (adult)	Hospital based (adult)	General population (adult)	Hospital based (mixed: children and adult)	Hospital based (adult)
Assessment of NAFLD	^{1}H-MRS	Liver biopsy	Liver biopsy	^{1}H-MRS	Liver biopsy	Controls: liver US and laboratory; patients: liver biopsy
Genotype frequencies (CC/CT/TT)	2,470/259/7	271/70/8	Unknown	794/122/4	1,044/155/2	304/ 52/ 5
Global study MAF	0.049	0.12* (MAF in patients)	Unknown	0.07	0.062	0.086
Association with NAFLD	$P = 5.7 \times 10^{-8}$	$P = 0.0008$*	Not explored	P = NS	Not explored	$P = 0.038$
Association with the degree of histological steatosis	Not explored	P = NS	P = NS	Not explored	$P = 0.001$	$P = 0.0299$
Association with NASH	Not explored	$P = 0.030$	P = NS	Not explored	$P = 0.003$ (OR, 1.84; 95% CI: 1.23-2.79)	$P = 0.027$ (OR, 1.66; 95% CI: 1.08-2.55)
Association with fibrosis	Not explored	$P = 5.57 \times 10^{-5}$ (OR, 2.94; 95% CI: 1.76-4.98)	$P = 0.014$ (OR, 1.46; 95% CI: 1.03-2.09)	P = NS (assessed by TE)	$P = 0.001$ (the effect was abolished after conditioning for NASH); not replicated in a bariatric surgery cohort	P = NS; (OR, 0.85; 95% CI: 0.60-1.20)
Association with ALT and AST	ALT‡: $P = 0.014$ AST: P = NS	Not explored	Not explored	ALT: P = NS AST: P = NS	ALT: P = NS AST: P = NS (dominant model)	ALT: P = NS AST: P = NS

Hepatic TG content was measured with ^{1}H-MRS.

*Association with NAFLD: In the absence of controls with phenotypic characterization, comparisons between genotype frequencies were done with 1000 Genomes European Caucasian population sample (n = 379); the samples for the 1000 Genomes Project mostly are anonymous and have no associated medical or phenotype data.

†Comparison of genotype frequencies in patients with NAFLD.

‡Association with ALT replicated in two large cohorts: Dallas Biobank (European Americans): $P = 0.003$; Copenhagen Study: $P = 7.6 \times 10^{-14}$.

Abbreviation: NS, nonsignificant.

Clinicians are aware that NAFLD patients have an increased risk of CVD[19,20]; in addition, NAFLD patients often show increased serum lipids levels, a risk factor for liver fat accumulation.[21]

Finally, the association of rs58542926 and the level of serum transaminases remains to be elucidated because the association was neither replicated consistently in different cohorts (Table 5) nor reported in previous large GWAS on liver enzymes.[22,23]

What Is the Functional Role of the TM6SF2-rs58542926 Genotypes on NAFLD Biology? The biological role of *TM6SF2* in lipid metabolism has been elucidated recently; for instance, Holmen et al. showed that *tm6sf2*-overexpressing mice have increased levels of lipids in circulation, including cholesterol, LDL, and TG, with no evidence of TG accumulation in the liver.[13] On the contrary, in experimental models, Kozlitina et al. demonstrated that inhibition of *tm6sf2* increased hepatocyte-TG content and decreased plasma levels of cholesterol, suggesting that the effect of rs58542926 on NAFLD results from a reduction in TM6SF2 function.[5] Accordingly, Mahdessian et al. showed that *TM6SF2* overexpression reduces liver cell steatosis, suggesting that *TM6SF2* is a regulator of liver fat metabolism with opposing effects on the secretion of TG-rich lipoproteins and hepatic lipid droplet content.[24]

It is known that TM6SF2 is expressed in the liver; thus, to understand whether the rs58542926 genotypes have any effect on human NAFLD, we explored the level of liver TM6SF2 expression in subjects with NAFLD at different stages of disease severity using IHC. Interestingly, we observed that TM6SF2 protein expression was significantly reduced in liver of patients with NAFLD; reduced expression of liver TM6SF2 was associated with a high degree of steatosis and NAS. In addition, we noted that liver TM6SF2 immunoreactivity was reduced in carriers of the NAFLD-risk T allele (Lys167); allelic-specific expression analysis of cDNA isolated from liver tissue confirmed that expression levels of rs58542926-T are approximately 56% of that of the C allele. Taken together, these findings suggest that the *TM6SF2*-NAFLD risk T allele is associated with decreased gene and protein expression in liver of affected patients. Further studies are required to investigate whether the variant affects gene transcription, mRNA, or protein half-life, as previously suggested.[5]

Consistent with previous *in vitro* findings demonstrating that TM6SF2 is localized in the endoplasmic reticulum (ER) and the ER-Golgi intermediate compartment of human liver cells,[24] we observed that TM6SF2 was expressed preferentially in the cytoplasm of hepatocytes. Overall, these observations suggest that TM6SF2 is involved in the biology of NAFLD; however, we cannot conclude that the reduced liver expression of TM6SF2 in NAFLD patients is a consequence or a cause of the disease.

In conclusion, rs58542926 is a low-frequency variant with a rather modest putative effect on NAFLD, suggesting that carriers of the risk T allele (Lys167) are slightly more likely to accumulate fat in the liver and develop NASH than those who are not, although the genetic model deserves more investigation. This observation supports the concept that other than "common-frequent" variants may contribute to the heritability of NAFLD and may also explain, at least in part, the missing heritability problem. Though the germinal finding from the exome-wide association study showed that the rs58542926 had a genome-wide significance for association with NAFLD,[5] the rs738409 located in the *PNPLA3* locus remains the most important contributor of the genetic risk of NAFLD. Notably, there would not seem to be any interaction between the effect of the *PNPLA3*−738409 and *TM6SF2*-rs58542926 variants on NAFLD. Although the complete genetic map of NAFLD is still being built, there is compelling evidence that liver fat accumulation is a multifactorial trait modulated by the effect of multiple genes,[2] including *PNPLA3* that not only regulates the morphology and physiology of liver lipid droplets,[25] but also global liver metabolism[26] and that *TM6SF2* is apparently required for normal very-low-density lipoprotein secretion.[5]

A challenge still remains in interpreting associations from GWAS and elucidating the causal effect of the reported associated NAFLD variants; for instance, is rs2228603 in the *NCAN* locus[27] or *TM6SF2*-rs58542926 responsible for the effect on fatty liver? Both SNPs are in high-linkage disequilibrium (pairwise r^2: 0.798; 1000genomes: phase_1_CEU).

Nevertheless, altogether, based on the results of previous association studies[5,6,17] and on our results showing that rs58542926 might regulate liver transcript and protein expression in an allele-specific manner, it is plausible to suggest that rs58542926 is the causal variant.

References

1. Loomba R, Sanyal AJ. The global NAFLD epidemic. Nat Rev Gastroenterol Hepatol 2013;10:686-690.
2. Sookoian S, Pirola CJ. The genetic epidemiology of nonalcoholic fatty liver disease: toward a personalized medicine. Clin Liver Dis 2012;16:467-485.
3. **Romeo S**, **Kozlitina J**, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461-1465.

4. 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;53:1883-1894.
5. **Kozlitina J**, **Smagris E**, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2014;46:352-356.
6. Liu YL, Reeves HL, Burt AD, Tiniakos D, McPherson S, Leathart JB, et al. TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. Nat Commun 2014;5:4309.
7. Wong VW, Wong GL, Tse CH, Chan HL. Prevalence of the TM6SF2 variant and non-alcoholic fatty liver disease in Chinese. J Hepatol 2014;61:708-709.
8. Sookoian S, Castano GO, Burgueno AL, Gianotti TF, Rosselli MS, Pirola CJ. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res 2009;50:2111-2116.
9. Pearson TA, Blair SN, Daniels SR, Eckel RH, Fair JM, Fortmann SP, et al. AHA Guidelines for Primary Prevention of Cardiovascular Disease and Stroke: 2002 Update: Consensus Panel Guide to Comprehensive Risk Reduction for Adult Patients Without Coronary or Other Atherosclerotic Vascular Diseases. American Heart Association Science Advisory and Coordinating Committee. Circulation 2002;106:388-391.
10. Kleiner DE, Brunt EM, Van NM, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. HEPATOLOGY 2005;41:1313-1321.
11. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. HEPATOLOGY 2011;53:810-820.
12. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155:945-959.
13. **Holmen OL**, **Zhang H**, Fan Y, Hovelson DH, Schmidt EM, Zhou W, et al. Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. Nat Genet 2014;46:345-351.
14. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 2006;38:209-213.
15. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet 2014;95:5-23.
16. Day CP. Is necroinflammation a prerequisite for fibrogenesis? Hepatogastroenterology 1996;43:104-120.
17. **Dongiovanni P**, **Petta S**, Maglio C, Fracanzani AL, Pipitone R, Mozzi E, et al. TM6SF2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease. HEPATOLOGY 2014 Sep 24. doi: 10.1002/hep.27490. [Epub ahead of print].
18. Chalasani N, Wilson L, Kleiner DE, Cummings OW, Brunt EM, Unalp A. Relationship of steatosis grade and zonal location to histological features of steatohepatitis in adult patients with non-alcoholic fatty liver disease. J Hepatol 2008;48:829-834.
19. Sookoian S, Pirola CJ. Non-alcoholic fatty liver disease is strongly associated with carotid atherosclerosis: a systematic review. J Hepatol 2008;49:600-607.
20. Lonardo A, Sookoian S, Chonchol M, Loria P, Targher G. Cardiovascular and systemic risk in nonalcoholic fatty liver disease—atherosclerosis as a major player in the natural course of NAFLD. Curr Pharm Des 2013;19:5177-5192.
21. Corey KE, Chalasani N. Management of dyslipidemia as a cardiovascular risk factor in individuals with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2014;12:1077-1084.
22. **Chambers JC**, **Zhang W**, **Sehmi J**, **Li X**, **Wass MN**, **van der Harst P**, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 2011;43:1131-1138.
23. Yuan X, Waterworth D, Perry JR, Lim N, Song K, Chambers JC, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet 2008;83:520-528.
24. Mahdessian H, Taxiarchis A, Popov S, Silveira A, Franco-Cereceda A, Hamsten A, et al. TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. Proc Natl Acad Sci U S A 2014;111:8913-8918.
25. Smagris E, BasuRay S, Li J, Huang Y, Lai KM, Gromada J, et al. Pnpla3I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis. HEPATOLOGY 2014 Jun 10. doi: 10.1002/hep.27242. [Epub ahead of print].
26. Min HK, Sookoian S, Pirola CJ, Cheng J, Mirshahi F, Sanyal AJ. Metabolic profiling reveals that PNPLA3 induces widespread effects on metabolism beyond triacylglycerol remodeling in Huh-7 hepatoma cells. Am J Physiol Gastrointest Liver Physiol 2014;307:G66-G76.
27. **Speliotes EK**, **Yerges-Armstrong LM**, **Wu J**, **Hernaez R**, **Kim LJ**, Palmer CD, 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.

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.27556/suppinfo.