

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

Silvia Sookoian¹ and Carlos J. Pirola²

Our objective was to estimate the strength of the effect of the I148M (rs738409 C/G) patatin-like phospholipase domain containing 3 (*PNPLA3*) variant on nonalcoholic fatty liver (NAFLD) and disease severity across different populations. We performed a systematic review by a meta-analysis; literature searches identified 16 studies. Our results showed that rs738409 exerted a strong influence not only on liver fat accumulation (GG homozygous showed 73% higher lipid fat content when compared with CC ones, data from 2,937 subjects; $P < 1 \times 10^{-9}$), but also on the susceptibility of a more aggressive disease (GG homozygous had 3.24-fold greater risk of higher necroinflammatory scores and 3.2-fold greater risk of developing fibrosis when compared with CC homozygous; $P < 1 \times 10^{-9}$; data from 1,739 and 2,251 individuals, respectively). Nonalcoholic steatohepatitis (NASH) was more frequently observed in GG than CC homozygous (odds ratio [OR] 3.488, 95% confidence interval [CI] 1.859-6.545, random model; $P < 2 \times 10^{-4}$; data from 2,124 patients). Evaluation of the risk associated with heterozygosity for the variant suggests that the additive genetic model best explains the effect of rs738409 on the susceptibility to develop NAFLD. Nevertheless, carrying two G alleles does not seem to increase the risk of severe histological features. Meta-regression showed a negative correlation between male sex and the effect of rs738409 on liver fat content (slope: -2.45 ± 1.04 ; $P < 0.02$). The rs738409 GG genotype versus the CC genotype was associated with a 28% increase in serum alanine aminotransferase levels. **Conclusion:** By summarizing the amount of evidence, this study provided unequivocal evidence of rs738409 as a strong modifier of the natural history of NAFLD in different populations around the world. (HEPATOLOGY 2011;53:1883-1894)

Advances in genome analysis, including high-throughput genotyping technology, have strongly contributed to the understanding of the genetic component of nonalcoholic fatty liver disease (NAFLD). In fact, the nonsynonymous variant I148M (rs738409 C/G, chr22:42656060-42656060) located in human patatin-like phospholipase domain

containing 3 gene (*PNPLA3*, also known as adiponutrin) was initially associated with fatty liver in the first genome-wide association study (GWAS) on NAFLD,¹ and further robustly replicated in independent candidate gene studies. The study of Romeo et al.,¹ a large multiethnic population-based epidemiological study of fatty liver, not only demonstrated that the rs738409

Abbreviations: ALT, alanine aminotransferase; GWAS, genome-wide association study; NAFLD, nonalcoholic fatty liver; NASH, nonalcoholic steatohepatitis; *PNPLA3*, patatin-like phospholipase domain containing 3.

From the ¹Department of Clinical and Molecular Hepatology; ²Department of Molecular Genetics and Biology of Complex Diseases, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires-National-Council of Scientific and Technological Research (CONICET), Ciudad Autónoma de Buenos Aires, Argentina.

Received December 29, 2010; accepted February 21, 2011.

Partially supported by grants PICT 2006-124 and PICT 2008-1521 (Agencia Nacional de Promoción Científica y Tecnológica), and UBACYT M55 (Universidad de Buenos Aires). The authors belong to Consejo Nacional de Investigaciones Científicas.

Address reprint requests to: Silvia Sookoian, M.D., Ph.D., or Carlos J. Pirola, Ph.D., Instituto de Investigaciones, Médicas A. Lanari-CONICET, Combatierte de Malvinas 3150, Buenos Aires (1427), Argentina. E-mail: ssookoian@lanari.fmed.uba.ar or pirola.carlos@lanari.fmed.uba.ar; fax: 54-11-4523-8947.

Copyright © 2011 by the American Association for the Study of Liver Diseases.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.24283

Potential conflict of interest: Nothing to report.

Additional Supporting Information may be found in the online version of this article.

variant is significantly associated with increased liver fat content, as evaluated by proton magnetic resonance spectroscopy, but also opened subsequent important questions about the pathogenesis and biology of NAFLD. The first question was about the potential association between the I148M variant and histological disease severity. Therefore, we first demonstrated,² a finding replicated by others,³⁻⁶ that the G allele in the forward strand was significantly associated with disease severity, as evaluated by histological assessment of liver biopsies. Nevertheless, despite the fact that nonalcoholic steatohepatitis (NASH) was more frequently observed among G allele carriers in some studies,^{2,4-6} others were unable to replicate this finding even after including a larger sample of cases.³ In addition, the magnitude and strength of the effect of the variant on disease severity widely varied among populations of similar ethnic background, ranging from odds ratios (ORs) of 1.5⁶ to 3.26⁴ for NASH, and ORs of 1.5⁶ to 3.37⁴ for liver fibrosis.

The concomitant question that came out after the findings of the GWAS was whether the rs738409 variant also influenced the prevalence and the behavior of NAFLD in early life. Therefore, large and well-characterized studies including pediatric population were carried out showing that some⁵ but not all of them³ confirmed an association between rs738409 and the severity of NAFLD in pediatric cohorts.

Almost simultaneously, the research community sought to know whether the metabolic comorbidities that are frequently associated with the pathogenesis of NAFLD, such as insulin resistance and obesity, are also influenced by rs738409. Interestingly, although most of the NAFLD studies have simultaneously confirmed the independence of the effect of rs738409 on fatty liver from the mentioned related phenotypes, a population-based study that surveyed a large sample of subjects without fatty liver ($n = 1,811$) has shown that *PNPLA3* variants, including rs738409, are associated with obesity and insulin sensitivity and secretion.⁷ In addition, the G allele of rs738409 is associated with a favorable metabolic profile, including decreased body mass index (BMI) and decreased risk of type 2 diabetes in one of the large NAFLD studies.⁴

The association of the I148M variant with increased liver enzymes, in particular alanine aminotransferase (ALT) levels, was first discovered by a GWAS of plasma liver-enzyme levels in three different populations,⁸ and thereafter replicated by several independent studies. In contrast to that observed in adults, the rs738409 variant was not associated with ALT levels in a series of pediatric patients with NAFLD, proven by liver biopsy.⁵

Finally, Romeo et al.¹ showed that the effect of the rs738409 variant was more pronounced among individuals of Hispanic ancestry, in whom the risk allele was also more frequent when compared with that of European-Americans and African-Americans. Hence, this observation had opened new investigations about the magnitude of the effect of the variant in different populations.

In view of the evidence mentioned above, our primary purpose was to estimate from the available literature the strength of the effect of the rs738409 variant on NAFLD and the histological disease severity across different populations, and the potential influence of the intermediate associated phenotypes. In addition, we systematically evaluated the study characteristics that could be responsible for the association. Furthermore, in order to provide novel information compared to what is already established in the literature, as the issue is still unresolved, we addressed in this meta-analysis which genetic model best explains the effect of the rs738409 single nucleotide polymorphism (SNP) on the susceptibility to develop NAFLD or NASH.

Materials and Methods

Data Sources and Study Selection. For the electronic searches, published studies were found through PubMed at the National Library of Medicine (<http://ncbi.nlm.nih.gov/entrez/query>) and in Medline databases for the query "(PNPLA3, adiponutrin) and (rs738409, gene or variants or polymorphism or alleles) and (fatty liver)." Reference lists in relevant publications were also examined. The literature search was done on studies up to January 2011 and availability of an English-language abstract or article for review; this yielded 32 hits. One study that performed a GWAS on NAFLD was excluded because rs738409 was not captured by the chip.⁹

There were no country restrictions. The authors reviewed all abstracts independently either to determine the eligibility criteria or for examining the appropriateness of the research issue and, when so, the article was retrieved; there were no discrepancies.

Details about inclusion and exclusion criteria and data collection can be seen in the Supporting Material online.

The evaluation of histological disease severity was based on data about liver biopsy of NAFLD patients, including the presence of NASH as defined by Kleiner et al.,¹⁰ presence of lobular necroinflammation (grade >1), and presence of fibrosis (stage >1).

Because the variation seemed to follow an undefined model of inheritance in some of the outcomes, to avoid choosing any *a priori* model, we decided to compare the extreme genotypes, namely, homozygous CC (148 I/I) versus homozygous GG (148 M/M), as reported.¹¹ In addition, in order to address which genetic model best explains the effect of the rs738409 SNP on the susceptibility to develop NAFLD and NASH, we also included an evaluation of the risk associated with heterozygosity for the variant (heterozygous CG versus homozygous CC, the reference group). For each phenotype we evaluated the association results stratified by age and ethnicity.

An evaluation of study quality of the reviewed articles using the median impact factor of the journals in which they had been published was included.¹²

Statistical Analysis. For quantitative variables, effect stands for standardized difference (D), defined as the mean difference (between GG and CC groups, and also between CG and CC groups) divided by the common within-group standard deviation, and for dichotomic variables, effect stands for OR with respect to the homozygous CC as a reference group unless indicated. Summary ORs and corresponding 95% CIs were estimated by fixed and random effects meta-analysis, respectively. Fixed and random effect models using the Mantel-Haenszel method were used to summarize results, obtaining the corresponding pooled OR.

For D, Cohen test (which is used for expressing the magnitude of differences between groups) was used to summarize the results, and heterogeneity was evaluated with Q statistic and the I^2 statistic, a transformation of Q that estimates the percentage of the variation in effect sizes that is due to heterogeneity. An I^2 value of 0% indicated no observed heterogeneity, and larger values showed an increasing heterogeneity. In the case of heterogeneity, we proceed as explained before¹³ (details can be seen in Supporting Material online). To check for publication bias, we used the Begg and Mazumdar's rank correlation test (this test reports the rank correlation, also known as rank correlation coefficient or simply Kendall's tau, between the standardized effect size and the variances, or standard errors, of these effects).¹⁴ A *P* value equal to or less than 0.05 was considered statistically significant.

All calculations were performed using the Comprehensive Meta-Analysis computer program (Biostat, Englewood, NJ).

Results

We evaluated 16 studies that met the selection criteria and that were identified using the search strategy

described in Supporting Fig. 1. Studies characteristics are shown in Table 1. Data from one study that fulfilled the eligibility criteria was included after personal contact with the investigators¹⁵; data on one further study was unavailable because in the article the authors did not disclose the raw data and our attempts to contact the authors were unsuccessful.¹⁶ All the studies scored well in terms of adequate descriptions of selection criteria and reference test, blind assessment of the reference test, and the availability of clinical data. A general critique concerns the observation that information about genotype counts per evaluated phenotype was scarcely found across the studies.

Study Characteristics. Eleven studies were hospital-based case-control studies,^{2-6,15,17-21} and the other five were population-based case-control studies,^{1,22-24} or family-based studies.²⁵

Information about liver biopsy was available in six studies,^{2-6,17} and data about disease severity was analyzed in 2,651 patients with NAFLD; ALT levels according to the rs738409 genotypes were available in 11 studies.^{1,2,5,6,15,17-21,24}

Genotyping for rs738409 was carried out across studies using Taqman assay in 11^{1,5,6,17-24} studies, by allele-specific oligonucleotides in two studies,^{2,15} and by Sequenom MassARRAY iPLEX Gold platform (Sequenom, San Diego, CA) in the remaining three studies.^{3,4,25}

Fatty Liver Disease and Liver Fat Content. Data regarding fatty liver disease as a disease trait extracted from 11 studies included 5,100 individuals,^{2-4,6,15,17,18,20,23-25} and, as expected, the analysis showed a significant association between fatty liver and the rs738409 variant either in the fixed or the random model ($P < 1 \times 10^{-9}$) (Fig. 1a); details of the association stratified by age are shown in Supporting Fig. 2. At any rate, we did not observe heterogeneity among studies as assessed by the Q statistic ($P = 0.33$), I^2 : 11.97. From the Begg and Mazumdar's rank correlation test (two-tailed $P = 0.15$), it seems that there was no publication bias. The evaluation of the risk associated with heterozygosity for the variant and fatty liver as a dichotomic variable also showed a significant association with the G allele. Interestingly, this analysis suggests that rs738409 exerts an additive effect on the susceptibility to develop NAFLD (Fig. 7); the details of the association analysis results for NAFLD and the CG versus CC genotypes are given in Supporting Table 1.

Table 1. Characteristics of the Studies on the Association Between the Nonsynonymous rs738409 Variant of *PNPLA3* and Fatty Liver Disease

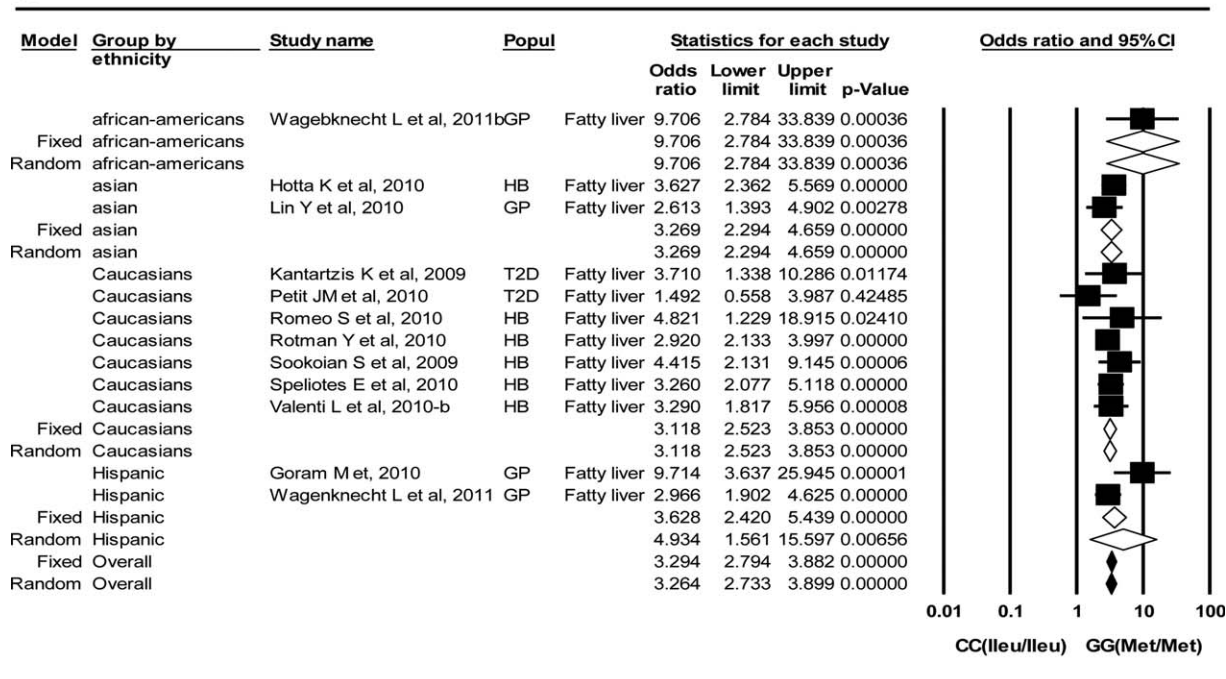
First Author, Year	Ref.	Population Ethnicity, Country	Study Design and Sample Size (N)	Features and Patients Characteristics	Liver Biopsy (N)	Age of the Subjects	Female Sex, n (%)
Romeo S, 2008	1	Mixed (Hispanics/ African-Americans/ Caucasian) USA	Population-based N=597/1814/1039	Hepatic steatosis measured by H-MRS	NA	Adult	Hispanics: 347 (58) African-Americans: 1049 (57.8) Caucasian: 543 (52.2)
Sookoian, 2009	2	European descendent (Caucasian) Argentina	Hospital-based N=266	Hepatic steatosis measured by US and LB	103	Adult	186 (69.9)
Kotronen, 2009	19	Finnish Finland	Hospital-based N=291	Hepatic steatosis measured by H-MRS	NA	Adult	160 (55)
Valenti, 2010	5	Caucasian Italy	Hospital-based N=149	Hepatic steatosis diagnosed by LB	149	Pediatric	56 (36.6)
Valenti, 2010	6	Caucasian Italy United Kingdom	Mixed: hospital-based and population-based. Italy: 432 UK: 321	Patients: Hepatic steatosis diagnosed by LB. Controls: normal ALT.	574	Adult	Italy: 114 (26.3) UK: 123 (38)
Romeo, 2010	21	Caucasian Italy	Hospital-based N=678	Hepatic steatosis: NA. Morbid obese	NA	Adult	459 (67.7)
Romeo, 2010	20	Caucasian Italy	Hospital-based N=475	Hepatic steatosis measured by US. Overweight/obese children	NA	Pediatric	248 (52.2)
Rotman, 2010	3	Caucasian USA	Mixed: population-based and hospital-based. N=894	Patients: Hepatic steatosis diagnosed by LB.	894	Adult	557 (62.3)
Speliotes, 2010	4	Caucasian USA	Mixed: population-based and hospital-based N=2128	Patients: Hepatic steatosis diagnosed by LB (n=678).	678	Adult	NA
Goran, 2010	23	Hispanic USA	Population-based N=327	Hepatic steatosis measured by DEXA	NA	Pediatric	129 (68.6)
Davis, 2010	22	Hispanic USA	Population-based N=153	Hepatic steatosis measured by H-MRS	NA	Pediatric	114 (74.5)
Petit, 2010		Caucasian France	Hospital-based N=218	Hepatic steatosis measured by H-MRS/ T2D	NA	Adult	111 (50.9)
Kantartzis, 2010	18	Caucasian Germany	Hospital-based N=330	Hepatic steatosis measured by H-MRS/ T2D	NA	Adult	200 (60.6)
Lin, 2010	24	Asian Taiwan	Population-based N=520	Overweight/obese children. Hepatic steatosis measured by US	NA	Pediatric	174 (33.5)
Hotta, 2010	17	Asian Japan	Hospital-based N=831	Patients: Hepatic steatosis diagnosed by LB. Controls: normal ALT.	253	Adult	527 (63.4)
Wagenknecht, 2011	25	Hispanic American African American USA	Population-based N=843/371	Hepatic steatosis measured by abdominal CT scanning	NA	Adult	Hispanic American 524 (62.6) African American 229 (61.7)

H-MRS: hydrogen magnetic resonance (H-MR) spectroscopy, US: liver ultrasonographic examination, CT: computed tomography, LB: liver biopsy, DEXA: dual energy x-ray absorptiometry, T2D: type 2 diabetes, NA: not available.

In addition, we found five homogeneous reports ($P = 0.22$, I^2 : 27.5) that reported retrieval data about the measurement of liver fat content (determined using hydrogen magnetic resonance spectroscopy [H-MRS]) according to the rs738409 genotypes.^{1,18,19,22,23} The comparison between cases and controls, including

2,937 subjects, showed that liver fat content was significantly associated with the GG genotype, either in random effect or fixed models ($P < 1 \times 10^{-9}$) (Fig. 1b), without evidence of publication bias (two-tailed $P = 0.37$) (details of the association stratified by ethnicity are shown in Supporting Fig. 3).

A



B

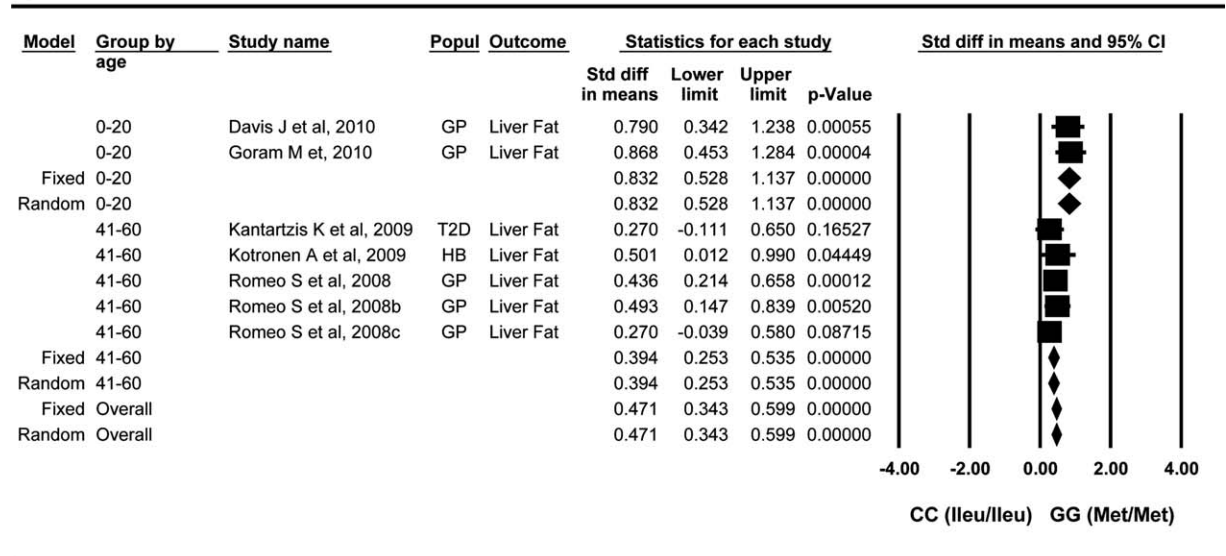


Fig. 1. (A) Summary estimates for ORs (effect), the corresponding 95% CI limits (lower and upper), and significance (P value) are calculated by fixed and random effects meta-analysis for fatty liver (as dichotomic variable) according to the rs738409 variant (homozygous CC and GG). The first author of the study and the year of publication are shown after citation name; Valenti L et al. 2010b stands for data about adult population. Wagebknecht L et al, 2011b stands for data about African American population. Studies are grouped by ethnicity. Popul: indicates design features, GP: general population, HB: hospital-based, and T2D: type 2 diabetes. In the graph, numbers indicate OR, filled squares stand for the effect of individual studies, and filled diamond express combined fixed and random effects. The symbol size is proportional to the number of individuals involved in each study. (B) Summary estimates for standardized difference (D) (effect), the corresponding 95% CI limits (lower and upper), and significance (P value) were estimated by fixed and random effects analysis for liver fat content (as continuous variable) according to the rs738409 variant (homozygous CC and GG). The first author of the study and the year of publication are shown after citation name; Romeo S et al. 2008 stands for Hispanics; Romeo S et al. 2008b stands for African-Americans; and Romeo S et al. 2008c for Caucasians. Popul: indicates design features, GP: general population, HB: hospital-based, and T2D: type 2 diabetes. In the graph, numbers indicate D values, filled squares stand for the effect of individual studies, and filled diamonds express combined fixed and random effects.

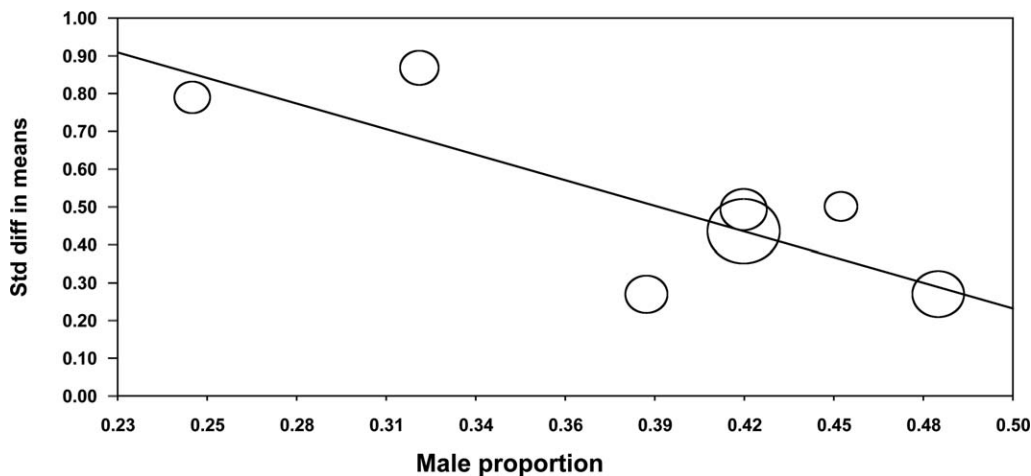


Fig. 2. Meta-regression analysis between male proportion among the studies and the effect of the rs738409 on liver fat content. Standardized difference (Std diff) for liver fat content is shown in Y-axis, and stands for the effect of the variant that is defined as the mean difference (between GG and CC groups) divided by the common within-group standard deviation. X-axis shows male proportion. Each circle represents a study with a size proportional to the sample size.

The evaluation of the risk associated with heterozygosity for the variant and liver fat content showed that, even significant, the effect seems to be much lower when carrying only one G allele (Fig. 7) (details in Supporting Table 1).

By meta-regression analysis, we observed a negative correlation between the male proportion in the studied populations and the effect of rs738409 on liver fat content (slope: -2.45 ± 1.04 , $P < 0.02$; Fig. 2), suggesting that a sexual dimorphism might be involved in the effect of the SNP on NAFLD development. Conversely, a significant correlation between the effect of the SNP on either NAFLD risk or liver fat content and BMI, and fasting glucose or fasting insulin could not be demonstrated (data not shown).

Histological Severity of NAFLD. We found six heterogeneous reports ($P < 0.001$, $I^2: 83.7$) that disclosed extractable data about the presence of NASH and either ORs per risk allele or the prevalence of NASH according to the rs738409 genotypes.^{2-6,17} The comparison among NAFLD patients, including 2,124 subjects with confirmed diagnosis by liver biopsy, showed that NASH was more frequently observed in GG than in CC carriers by fixed (3.125, 95% CI 2.690-3.630; $P < 1 \times 10^{-9}$) or random effect (3.488, 95% CI 1.859-6.545; $P < 2 \times 10^{-4}$) models, without evidence of publication bias (two-tailed $P = 0.45$); details of the association stratified by ethnicity are shown in Supporting Fig. 4. To investigate the source of heterogeneity, we analyzed the data by grouping the reports by age, and after separating one study that included a pediatric population and showed a disparate high OR of 88.6⁵ (Fig. 3), the heterogeneity still persisted

between the remaining four studies that included an adult population. The heterogeneity disappeared after excluding one outlier study,³ and the effect was still significant (OR 3.223, 95% CI 2.849-3.875, fixed and random model; $P < 1 \times 10^{-9}$).

Data about lobular necroinflammation according to either genotypes or ORs per risk allele was available in four heterogeneous studies ($P < 0.002$, $I^2: 79$),²⁻⁵ including 1,739 patients. The analysis showed that the GG genotype was significantly associated with higher inflammation scores (fixed $P < 1 \times 10^{-9}$ and random $P < 1 \times 10^{-7}$), without evidence of publication bias (two-tailed $P = 0.31$; Fig. 4). By separating one report⁵ that included pediatric patients (and again showed a disparate high OR of 72) the heterogeneity was removed, and the effect was still significant (OR 3.18, 95% CI 2.77-3.64, fixed and random model; $P < 1 \times 10^{-9}$).

Finally, data about fibrosis score was extractable from five homogeneous studies,²⁻⁶ including 2,251 patients. We observed a significant association with the variant either in fixed or random model ($P < 1 \times 10^{-9}$); there was no evidence of publication bias (two-tailed $P = 0.81$; Fig. 5). Owing to the limited number of studies that disclosed data about the NAFLD traits associated with disease severity and metabolic syndrome intermediate phenotypes, meta-regression analysis of these variables with BMI, homeostasis model assessment of insulin resistance (HOMA-IR), fasting glucose, or fasting insulin was not possible.

The evaluation of the risk associated with heterozygosity for the variant and liver disease severity showed that the effect when carrying only one G allele does not differ from the GG genotype (Fig. 7),

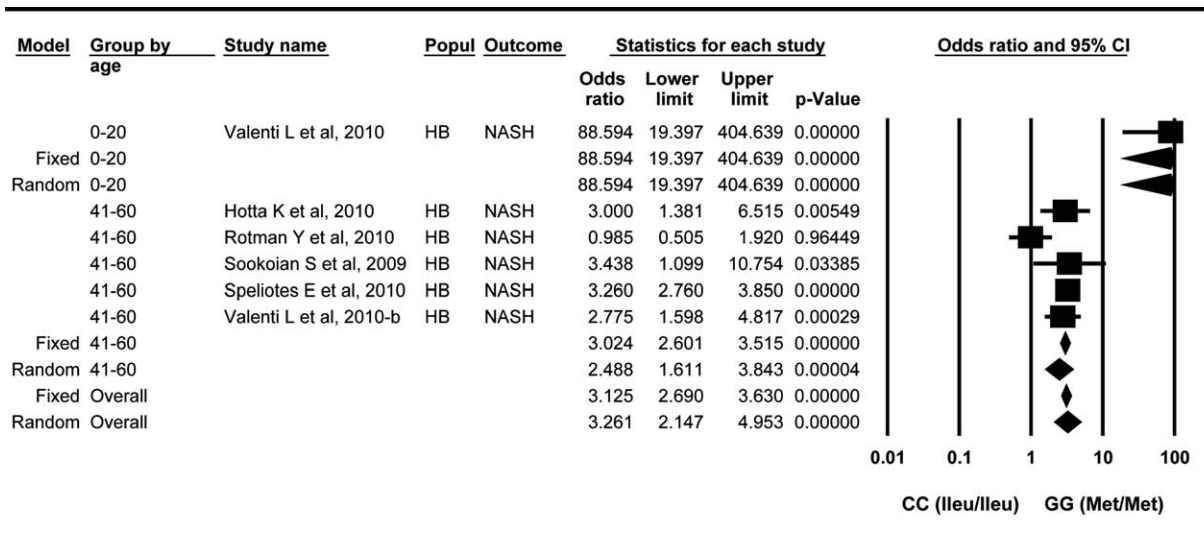


Fig. 3. Summary estimates for ORs (effect), the corresponding 95% CI limits (lower and upper) and significance (P value) were calculated by fixed and random effects meta-analysis for NASH according to the rs738409 variant (homozygous CC and GG). The first author of the study and the year of publication are shown after citation name; Valenti L et al. 2010b stands for data about adult population. Popul: indicates design features, GP: general population, HB: hospital-based, and T2D: type 2 diabetes. In the graph, numbers indicate OR, filled squares stand for the effect of individual studies, and filled diamonds express combined fixed and random effects. The symbol size is proportional to the number of individuals involved in each study.

suggesting that carrying two G alleles does not lead to a large change on the risk of severe histological features (details in Supporting Table 1).

ALT Levels. ALT was significantly associated with the rs738409 variant in 11 heterogeneous studies ($P < 0.001$, $I^2: 86.5$),^{1,2,5,6,15,17-21,24} including 5,366 individuals; fixed effect $P < 1 \times 10^{-9}$, and random effect $P < 0.0009$, without evidence of publication bias (two-tailed $P = 0.30$); details of the association stratified by

ethnicity are shown in Supporting Fig. 5. Subjects were stratified by age (Fig. 6), ethnicity, study design, and associated disease condition, but the heterogeneity remained significant. The heterogeneity did not disappear even after removing the outlier studies.^{5,18,19,21} Nevertheless, the effect estimate seems to be robust because similar and significant results (standard deviation between 0.32 and 0.45, $P < 1 \times 10^{-8}$) still remained after excluding one study at a time.

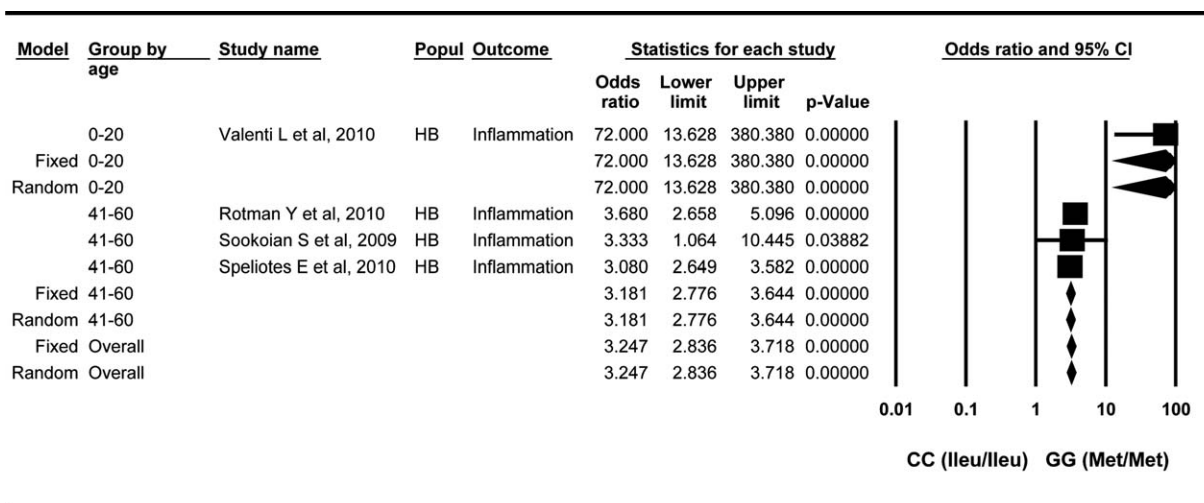


Fig. 4. Summary estimates for ORs (effect), the corresponding 95% CI limits (lower and upper) and significance (P value) were calculated by fixed and random effects meta-analysis for liver necroinflammation and the rs738409 variant (homozygous CC and GG). The first author of the study and the year of publication are shown after citation name. Popul: indicates design features, GP: general population, HB: hospital-based, and T2D: type 2 diabetes. In the graph, numbers indicate OR, filled squares stand for the effect of individual studies, and filled diamonds express combined fixed and random effects. The symbol size is proportional to the number of individuals involved in each study.

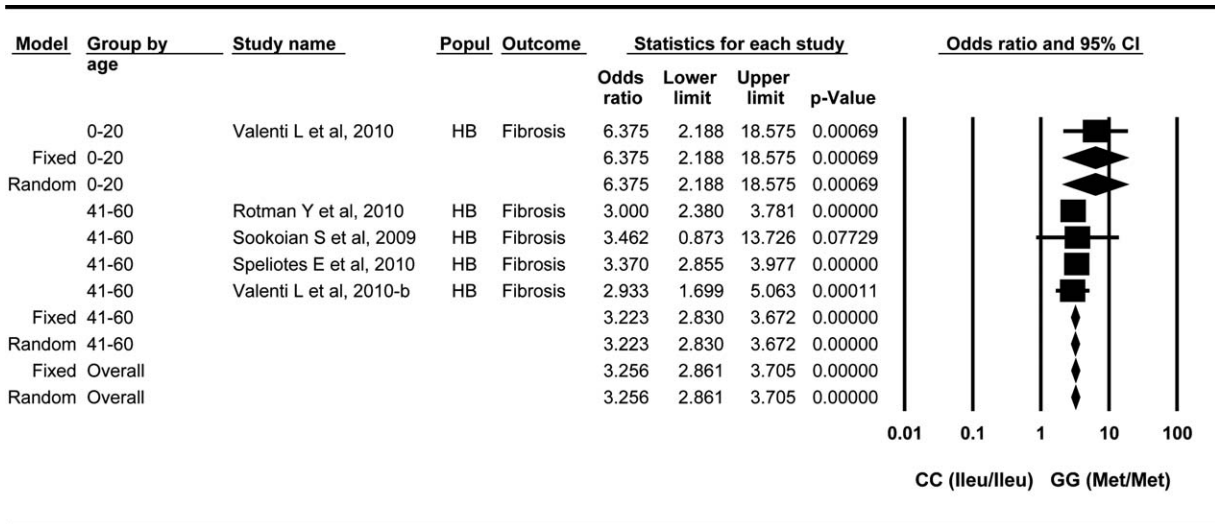


Fig. 5. Summary estimates for ORs (effect), the corresponding 95% CI limits (lower and upper) and significance (P value) were calculated by fixed and random effects meta-analysis for liver fibrosis and the rs738409 variant (homozygous CC and GG). The first author of the study and the year of publication are shown after citation name; Valenti L et al. 2010b stands for data about adult population. Popul: indicates design features, GP: general population, HB: hospital-based, and T2D: type 2 diabetes. In the graph, numbers indicate OR, filled squares stand for the effect of individual studies, and filled diamonds express combined fixed and random effects. The symbol size is proportional to the number of individuals involved in each study.

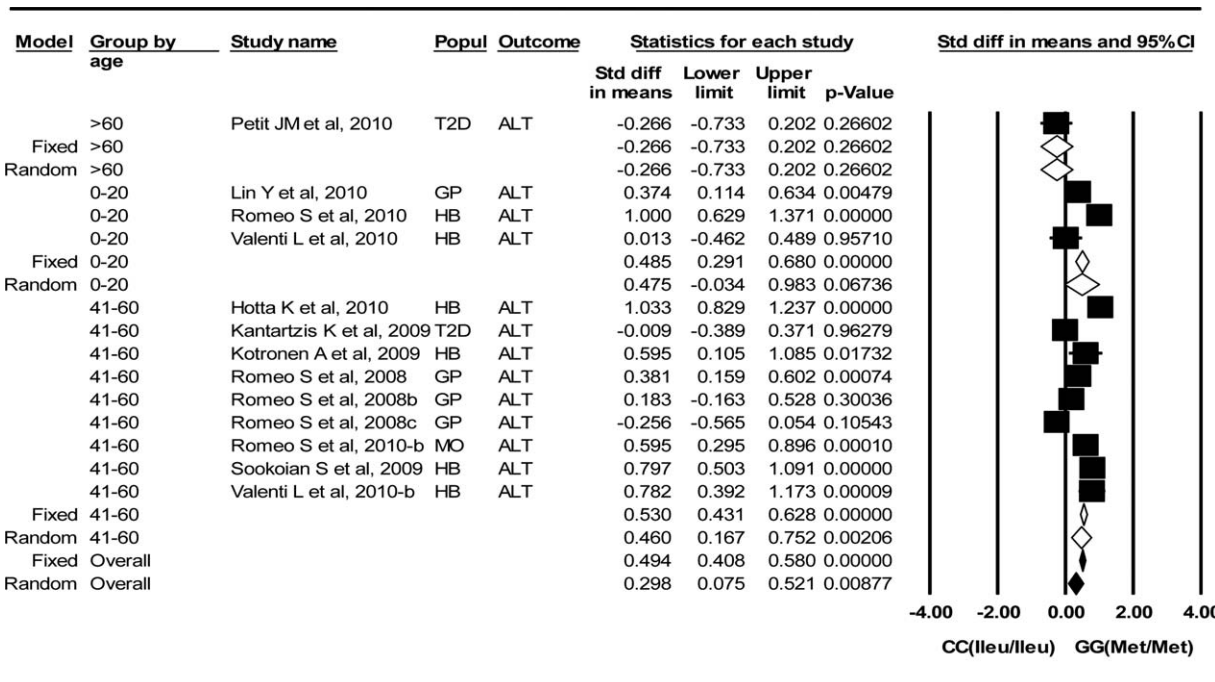


Fig. 6. Summary estimates for standardized difference (D) (effect), the corresponding 95% CI limits (lower and upper) and significance (P value) were estimated by fixed and random effects analysis for ALT (as continuous variable) according to the rs738409 variant (homozygous CC and GG). The first author of the study and the year of publication are shown after citation name; when the first author of the study appears more than once in the figure under the same year of publication, b and c stand for data either extracted from different populations or different citations (Valenti L et al. 2010b stands for data about adult population; Romeo S et al. 2008 stands for Hispanics; Romeo S et al. 2008b stands for African-Americans; Romeo S et al. 2008c for Caucasians; and Romeo S 2010b stands for adults with morbid obesity). Popul: indicates design features, GP: general population, HB: hospital-based, MO: morbid obesity, and T2D: type 2 diabetes. In the graph, numbers indicate D values, filled squares stand for the effect of individual studies, and filled diamonds express combined fixed and random effects.

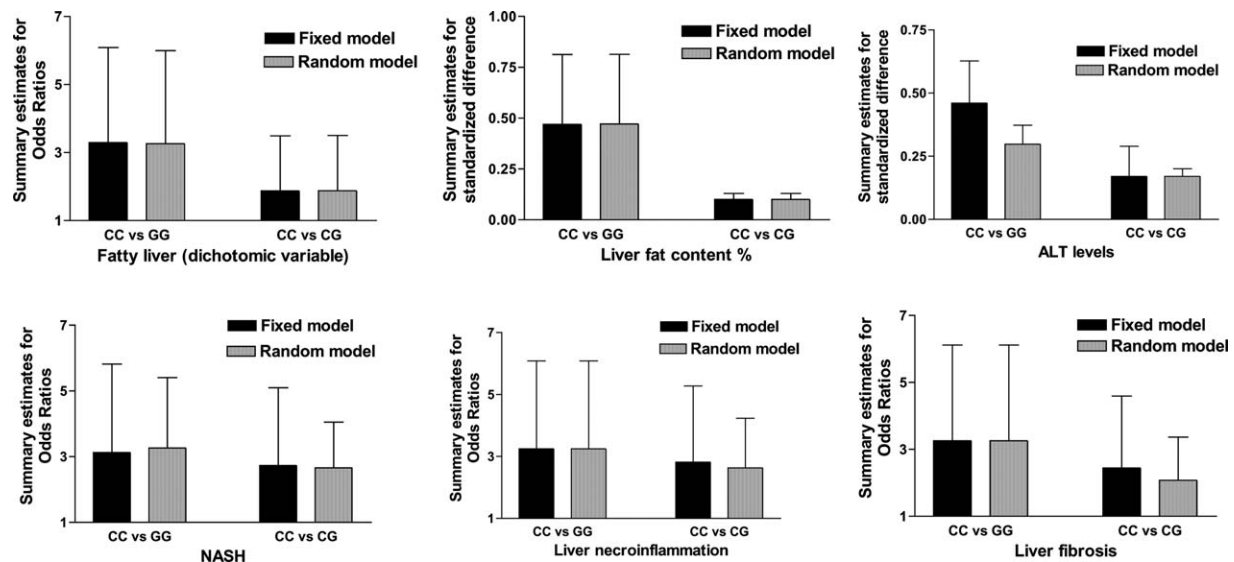


Fig. 7. Comparison of the magnitude of the effect of different genetic models for the rs738409 variant on the susceptibility to develop NAFLD and NASH-related phenotypes.

The analysis of the heterozygosity for the variant showed that ALT levels were significantly associated with the rs738409 G allele when the reference genotype (CC) was compared with the CG genotype, suggesting again an additive genetic effect (details in Supporting Table 1).

Associated Metabolic Phenotypes: Insulin Resistance and BMI. Additional information about the NAFLD-associated insulin resistance phenotype was available in six studies that reported data on HOMA-IR,^{2,5,6,20,21,24} including 1,404 subjects; in seven studies that reported data on fasting insulin levels,^{2,6,18-21,24} including 1,721 subjects; and in nine studies that reported data on fasting glucose levels,^{1,2,5,6,18-21,24} including 4160 subjects (Supporting Table 2). Interestingly, no significant association with the variant was found for HOMA-IR and fasting glucose or insulin levels. Neither was the trait obesity, as measured by BMI, associated with the variant (Supporting Table 2) in 4,141 subjects included in 10 heterogeneous studies.^{1,2,6,15,18,19,21,22}

Overall Study Quality. The median impact factor for all the included studies was 7.81 (range, 2.84-34.28). In conclusion, we observed that the data we have included in the analysis were published in leading journals with a high impact factor.

Discussion

Whereas the physiological role and biological function of *PNPLA3* in the liver is still unclear, the evi-

dence on the impact of the rs738409 nonsynonymous I148M polymorphism not only on the hepatic triglyceride content but on a more severe course of NAFLD seems to be irrevocable. The results of this well-powered meta-analysis, by summarizing the amount of evidence, degree of replication, and absence of publication bias, show that rs738409 exerts a strong influence not only on liver fat accumulation (GG carriers have a 73% higher lipid fat content when compared with CC carriers), but also on the susceptibility of a more aggressive disease with higher liver injury (GG homozygous subjects have 3.24-fold more risk of higher necroinflammatory scores when compared with those homozygous for the C allele; data from 1,739 individuals) and fibrosis scores (GG homozygous subjects have 3.2-fold more risk of develop liver fibrosis when compared with those homozygous for the C allele; data from 2,251 patients). These results coupled data from studies obtaining the largest available series of patients with NAFLD proven by liver biopsy.

In addition, novel information is added as we observed that the effect of rs738409 on fatty liver disease seems to follow an additive genetic effect for all phenotypes except for liver disease severity. This observation suggests that the influence of the variant on the susceptibility to develop a more progressive disease follows a dominant model.

Notably, meta-regression showed a negative association between the effect of rs738409 on liver fat content and male sex, novel data never analyzed before even in the large studies. This effect may explain the

previously reported sex differences in the prevalence of NAFLD²⁶ and, even more interesting, may suggest a potential sexual dimorphism in the effect of the gene variant on NAFLD susceptibility. At any rate, the different proportion of males in different studies explained, at least in part, the variation in the estimated effect of the gene variant on liver fat content. Despite the explanation on the gene by sex interaction cannot be given by the present study, some evidence exists that the *PNLPA3* gene is under the control of several factors, for instance, nutritional control in close relationship with *SREBP-1c* and liver X receptor.²⁷ Interestingly, sexual hormones, like estrogen, modulate lipogenic genes, including *SREBP-1c*, and participate in adiposity and fuel partitioning.²⁸ Hence, a possible mechanism could be the sex hormones modulation.

A note of caution should be added, as the presence of heterogeneity may potentially restrict the interpretation of the pooled risk estimates, in particular concerning the association of the variant with closely related features such as the presence of NASH, scores of liver necroinflammation, and serum ALT levels. Heterogeneity in a meta-analysis is mostly produced by differences in study design and background characteristics of the subjects, and the extent of heterogeneity might influence the conclusions. However, the random effect model, where heterogeneity is no longer a main issue, provided a significant result about these features closely related with the disease severity and the magnitude of the liver injury. Although heterogeneity was addressed statistically by applying a random effect model, we aimed to further investigate its potential sources where possible. Thus, the full dataset was utilized for investigation of heterogeneity by sensitivity analysis. In some studies, NAFLD patients were mostly recruited because of elevated ALT levels; thus, there was a marked enrichment of patients with NASH and few of them showed simple steatosis. Moreover, potential selection bias when selecting patients for liver biopsy may also explain the heterogeneity.

The current meta-analysis is useful to clearly understand the magnitude of the effect of rs738409 on the histological severity of NAFLD, which is far beyond the small magnitude observed for common variants on complex traits,²⁹ and may be explained by the nonsynonymous nature of the polymorphism that induces an amino acid change of I for M (missense Ileu (ATC) → Met (ATG)) with possible functional consequences.³⁰

A potential limitation of this study is its reliance on two studies,^{3,4} which included cases from the Nonalcoholic Steatohepatitis Clinical Research Network

(NASH CRN). This fact may limit the research results to liver disease severity. Hence, to circumvent this potential caveat, we performed the whole analysis about all the NAFLD-related phenotypes excluding the smaller study in terms of sample size³ (details in Supporting Table 3). We observed a very similar magnitude in the variant effect on the analyzed phenotypes.

Data about a GWAS on NAFLD performed in female adults also from the NASH CRN was not included in this meta-analysis because rs738409 was not captured by the chip.⁹ Although 15 SNPs in the *PNPLA3* were captured by the GWAS and none of them showed significant association with NAFLD, it is important to note that among five variants, only rs2076211 was in moderate linkage disequilibrium (r^2 : 0.65) with rs738409 precluding any imputation even assuming access to the complete dataset.

An interesting observation about the analysis of the significant association between rs738409 and liver enzymes is the 28% increase in serum levels observed in GG homozygous individuals, a relevant aspect that might be regarded when selecting patients for clinical trials, decision-making on indication of liver biopsy, and evaluation of the magnitude of any treatment response based on serum ALT levels.

The question of whether some NAFLD-associated metabolic comorbidities such as insulin resistance and obesity are also influenced by the rs738409 variant can also be answered by this meta-analysis, because all the included studies in which these phenotypes were measured showed a lack of significant difference among genotypes for HOMA-IR (data from 1,404 individuals), glucose and insulin levels (data from 4,160 and 1,721 subjects, respectively), and BMI (data from 4,141 individuals). These results, which are hardly due to lack of statistical power, are in agreement with data from large population-based association studies from several consortia that evaluated subjects without the associated NAFLD phenotype.⁴

As a final observation, we strongly encourage the inclusion in genetic association studies of all the data in standard format (complete record of all of the phenotypes available by genotypes and genotype counts for cases and controls), either in the print or supplementary electronic material, to facilitate further meta-analysis with more robust estimates of the genetic effect.

The Human Genome Epidemiology Network recommends that a meta-analysis is necessary before the evidence for a particular association can be regarded as strong.³¹ Therefore, we provide strong statistical evidence for the impact of the rs738409 variant on the

clinical course of NAFLD, including susceptibility to liver injury and fibrosis progression. A conservative population attributable risk (%) estimated for the GG genotype may be as high as 33 in Hispanics with a minimum allele frequency (MAF) of around 50% to as low as 9 in African-Americans who have an MAF below 25%. Ethnic differences in susceptibility to NAFLD are evident, as shown in previous studies that described a high prevalence in Hispanics, and significantly lower in African-Americans,³² but probably reflect differences in the variant MAF and not a different risk associated with the variant. Considering that NAFLD has also reached epidemic proportions in China and Japan,³³ the strength of this study is the evaluation of the effect of the variant on NAFLD in Asians, which compared with Caucasians is almost identical (3.26 versus 3.11) (Fig. 1a).

The future challenge of the medical community will be to manage the expectations about rapidly translating this knowledge into more individualized decision-making and personalized medicine. Nevertheless, the observed effect of the rs738409 variant on the behavior of NAFLD is perhaps one of the strongest ever reported for a common variant modifying the genetic susceptibility for complex diseases.

References

- 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.
- Sookoian S, Castano GO, Burgueno AL, Fernandez GT, Rosselli MS, Pirola CJ. A nonsynonymous gene variant in adiponutrin gene is associated with nonalcoholic fatty liver disease severity. *J Lipid Res* 2009;50:2111-2116.
- Rotman Y, Koh C, Zmuda JM, Kleiner DE, Liang TJ. The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease. *HEPATOLOGY* 2010;52:894-903.
- Speliotes EK, Butler JL, Palmer CD, Voight BF, Hirschhorn JN. PNPLA3 variants specifically confer increased risk for histologic nonalcoholic fatty liver disease but not metabolic disease. *HEPATOLOGY* 2010;52:904-912.
- Valenti L, Alisi A, Galmozzi E, Bartuli A, Del MB, Alterio A, et al. I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. *HEPATOLOGY* 2010;52:1274-1280.
- Valenti L, Al-Serri A, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, et al. Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. *HEPATOLOGY* 2010;51:1209-1217.
- Johansson LE, Lindblad U, Larsson CA, Rastam L, Ridderstrale M. Polymorphisms in the adiponutrin gene are associated with increased insulin secretion and obesity. *Eur J Endocrinol* 2008;159:577-583.
- 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.
- Chalasan N, Guo X, Loomba R, Goodarzi MO, Haritunians T, Kwon S, et al. Genome-wide association study identifies variants associated with histologic features of nonalcoholic fatty liver disease. *Gastroenterology* 2010;139:1567-76, 1576.
- 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.
- Sookoian S, Gianotti TF, Gonzalez CD, Pirola CJ. Association of the C-344T aldosterone synthase gene variant with essential hypertension: a meta-analysis. *J Hypertens* 2007;25:5-13.
- Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *HEPATOLOGY* 2006;44:865-873.
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;33:177-182.
- Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088-1101.
- Petit JM, Guieu B, Masson D, Duvallard L, Jooste V, Buffier P, et al. Specifically PNPLA3-mediated accumulation of liver fat in obese patients with type 2 diabetes. *J Clin Endocrinol Metab* 2010;95:E430-E436.
- Santoro N, Kursawe R, D'Adamo E, Dykas DJ, Zhang CK, Bale AE, et al. A common variant in the patatin-like phospholipase 3 gene (PNPLA3) is associated with fatty liver disease in obese children and adolescents. *HEPATOLOGY* 2010;52:1281-1290.
- Hotta K, Yoneda M, Hyogo H, Ochi H, Mizusawa S, Ueno T, et al. Association of the rs738409 polymorphism in PNPLA3 with liver damage and the development of nonalcoholic fatty liver disease. *BMC Med Genet* 2010;11:172.
- Kantartzis K, Peter A, Machicao F, Machann J, Wagner S, Konigsrainer I, et al. Dissociation between fatty liver and insulin resistance in humans carrying a variant of the patatin-like phospholipase 3 gene. *Diabetes* 2009;58:2616-2623.
- Kotroń A, Johansson LE, Johansson LM, Roos C, Westerbacka J, Hamsten A, et al. A common variant in PNPLA3, which encodes adiponutrin, is associated with liver fat content in humans. *Diabetologia* 2009;52:1056-1060.
- Romeo S, Sentinelli F, Cambuli VM, Incani M, Congiu T, Matta V, et al. The I48M allele of the PNPLA3 gene is associated with indices of liver damage early in life. *J Hepatol* 2010;53:335-338.
- Romeo S, Sentinelli F, Dash S, Yeo GS, Savage DB, Leonetti F, et al. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. *Int J Obes (Lond)* 2010;34:190-194.
- Davis JN, Le KA, Walker RW, Vikman S, Spruijt-Metz D, Weigensberg MJ, et al. Increased hepatic fat in overweight Hispanic youth influenced by interaction between genetic variation in PNPLA3 and high dietary carbohydrate and sugar consumption. *Am J Clin Nutr* 2010;92:1522-1527.
- Goran MI, Walker R, Le KA, Mahurkar S, Vikman S, Davis JN, et al. Effects of PNPLA3 on liver fat and metabolic profile in Hispanic children and adolescents. *Diabetes* 2010;59:3127-3130.
- Lin YC, Chang PF, Hu FC, Yang WS, Chang MH, Ni YH. A common variant in the PNPLA3 gene is a risk factor for non-alcoholic fatty liver disease in obese Taiwanese children. *J Pediatr*. December 17, 2010 [Epub ahead of print].
- Wagenknecht LE, Palmer ND, Bowden DW, Rotter JI, Norris JM, Ziegler J, et al. Association of PNPLA3 with non-alcoholic fatty liver disease in a minority cohort: the Insulin Resistance Atherosclerosis Family Study. *Liver Int* 2011;31:412-416.
- Sanyal AJ. AGA technical review on nonalcoholic fatty liver disease. *Gastroenterology* 2002;123:1705-1725.
- Huang Y, He S, Li JZ, Seo YK, Osborne TF, Cohen JC, et al. A feed-forward loop amplifies nutritional regulation of PNPLA3. *Proc Natl Acad Sci U S A* 2010;107:7892-7897.

28. D'Eon TM, Souza SC, Aronovitz M, Obin MS, Fried SK, Greenberg AS. Estrogen regulation of adiposity and fuel partitioning. Evidence of genomic and non-genomic regulation of lipogenic and oxidative pathways. *J Biol Chem* 2005;280:35983-35991.
29. Sookoian S, Pirola CJ. Metabolic syndrome: from the genetics to the pathophysiology. *Curr Hypertens Rep* 2011;13:149-157.
30. He S, McPhaul C, Li JZ, Garuti R, Kinch L, Grishin NV, et al. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem* 2010;285:6706-6715.
31. Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;37:120-132.
32. Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *HEPATOLOGY* 2004;40:1387-1395.
33. Angulo P. GI epidemiology: nonalcoholic fatty liver disease. *Aliment Pharmacol Ther* 2007;25:883-889.