Relationship between *PNPLA3* rs738409 polymorphism and decreased kidney function in children with NAFLD

Short title: PNPLA3 and kidney function

Giovanni Targher, MD[1], Alessandro Mantovani, MD[1], Anna Alisi, PhD[2], Antonella Mosca, MD[3], Nadia Panera, PhD[2], Christopher D. Byrne, MB BCh, PhD[4,5], Valerio Nobili, MD[3,6]

[1]Section of Endocrinology, Diabetes and Metabolism, Department of Medicine, University and Azienda Ospedaliera Universitaria Integrata of Verona, Verona, Italy

[2]Research Unit of Molecular Genetics of Complex Phenotypes, IRCCS "Bambino Gesù" Children's Hospital, Rome, Italy

[3]Hepatology, Gastroenterology and Nutrition Unit, IRCCS "Bambino Gesù" Children's Hospital, Rome, Italy

[4]Nutrition and Metabolism, Faculty of Medicine, University of Southampton, Southampton, UK

[5]Southampton National Institute for Health Research Biomedical Research Centre, University Hospital Southampton, Southampton General Hospital, Southampton SO16 6YD, UK

[6]Department of Pediatric, University "La Sapienza", Rome, Italy

Word count: abstract 255; text 5,166 (*including* abstract, text and references); n. 5 **Tables** + n. 2 **Figures**

Keywords: adiponutrin; kidney dysfunction; NAFLD; pediatrics

Address for correspondence:
Prof. Valerio Nobili, MD
Hepatology, Gastroenterology and Nutrition Unit
IRCCS "Bambino Gesù" Children's Hospital
Piazza Sant' Onofrio, 4
00165 Rome, Italy
E-mail: nobili66@yahoo.it

ABBREVIATION LIST

PNPLA3, patatin-like phospholipase domain-containing protein 3

NAFLD, nonalcoholic fatty liver disease

e-GFR, estimated glomerular filtration rate

NASH, nonalcoholic steatohepatitis

CKD, chronic kidney disease

BMI, body mass index

AST, aspartate aminotransferase

ALT, alanine aminotransferase

GGT, gamma-glutamyltransferase

HOMA, homeostasis model assessment

CRP, C-reactive protein

IL-6, interleukin-6

IFG, impaired fasting glycemia

NAS, NAFLD activity score

TNF, tumor necrosis factor

Disclosure Statement: The Authors have no potential conflicts of interest to disclose.

Funding Sources: GT is supported in part by grants from the University School of Medicine of Verona, Verona, Italy. CDB is supported in part by grants from the Southampton National Institute for Health Research Biomedical Research Centre.

Authors' Contributions: GT and VN conceived and designed the study. AA, AM, NP and VN researched data and reviewed/edited the manuscript. AM and CDB contributed to discussion and reviewed/edited the manuscript. GT and AM analyzed the data. GT wrote the manuscript. VN is the guarantor of this work and, as such, had full access to all the data of the study and take responsibility for the integrity and accuracy of data. All authors approved the final version of the manuscript.

ABSTRACT

Emerging evidence suggests that *patatin-like phospholipase domain-containing protein-3* (*PNPLA3*) rs738409 genotype (the major genetic variant associated with susceptibility to non-alcoholic fatty liver disease [NAFLD]) is associated with decreased kidney function in adults. Currently, it is uncertain whether this association also occurs in children/adolescents and whether any association is independent of liver disease severity. We enrolled a sample of 142 consecutive Caucasian children and adolescents with biopsy-proven NAFLD, presenting to the Liver Unit of the "Bambino Gesù" Children's Hospital. Glomerular filtration rate (e-GFR) was estimated using the Bedside Schwartz equation, whereas 24-hour proteinuria was measured with a radioimmunoassay method. Genotyping for *PNPLA3* rs738409 genotype was undertaken using the TaqMan SNP genotyping allelic discrimination method. Overall, forty-five children had G/G, 56 had G/C and 41 had C/C *PNPLA3* rs738409 genotype, respectively. Children with G/G genotype had significantly lower e-GFR (107.5±20 *vs*. 112.8±18 *vs*. 125.3±23 ml/min/1.73 m^2, p=0.002) and higher 24-hour proteinuria (58.5±21 *vs*. 53.9±22 *vs*. 42.9±20 mg/day, p=0.012) compared to those with either G/C or C/C genotypes. After adjustment for age, sex, measures of adiposity, systolic blood pressure, HOMA-estimated insulin resistance and biopsy-confirmed non-alcoholic steatohepatitis (NASH) and stage of fibrosis, the presence of rs738409 G/G genotype was independently associated with both lower e-GFR (β coefficient: -23.6, 95% CI -36.3 to -10.8, p<0.001) and higher 24-hour proteinuria (β coefficient: 15.3, 95% CI 1.12 to 30.5, p=0.046). **Conclusion:** Regardless of established renal risk factors and presence of NASH, the *PNPLA3* G/G genotype was strongly associated with decreasing kidney function and increasing 24-hour proteinuria in children/adolescents with histologically confirmed NAFLD.

INTRODUCTION

Pediatric non-alcoholic fatty liver disease (NAFLD) has become a major public health problem due to the striking increase in its prevalence and the clinical implications for future development of liver dysfunction, type 2 diabetes and other cardiometabolic complications (1-3). Indeed, NAFLD is the most common cause of chronic liver disease in children and adolescents in many parts of the world (1-3), and increasing evidence in adults now also suggests that NAFLD adversely affects kidney function and may be causally involved in the development of chronic kidney disease (CKD) (4,5).

To date, it is uncertain whether more severe liver disease, within the spectrum that encompasses NAFLD, is associated with a greater deterioration of kidney function, especially in the pediatric population. Among the genetic factors that may influence the progression of NAFLD, the minor allele G of rs738409, *i.e*., a non-synonymous single nucleotide polymorphism in the *patatin-like phospholipase domain-containing protein 3* (*PNPLA3*) gene encoding an Ile148Met change, has been recognized to be the major common genetic variant associated with a greater predisposition to progressive forms of NAFLD (6,7). The *PNPLA3* gene is mostly expressed in the liver and has an acyl hydrolase activity. In particular, the G allele of rs738409 is associated with the loss of hydrolyzing function of the protein, thus resulting in accumulation of lipid droplets into the hepatocytes (6,7). Some recent studies have also suggested that the G allele of rs738409 is significantly associated with lower values of estimated glomerular filtration rate (e-GFR) and a higher prevalence of CKD in middle-aged and elderly individuals, irrespective of the coexistence of NAFLD (8,9).

Presently, it is uncertain whether a significant association between *PNPLA3* rs738409 polymorphism and decreasing kidney function also occurs in the pediatric population. As children are less susceptible to developing hypertension or type 2 diabetes (two potential confounding diseases that are well known to adversely affect kidney function), we reasoned that testing the association between the *PNPLA3* rs738409 polymorphism and renal function in children would better elucidate associations between liver disease severity in NAFLD and abnormal kidney function. Therefore, the major aim of this cross-sectional study was to examine whether, and to what extent, *PNPLA3* rs738409 polymorphism is associated with decreasing kidney function and increasing 24-hour urinary protein excretion in children with biopsy-proven NAFLD.

METHODS

Subjects

In this exploratory analysis, we enrolled 142 unrelated children and adolescents (5–17 years) with an ultrasonographic diagnosis of severe hepatic steatosis or persistently (≥6 months) elevated serum aminotransferase levels, who were referred to the Liver Unit of the "Bambino Gesù" Children's Hospital between January 2013 and December 2015, and who accepted to undergo liver biopsy for confirming the presence and severity of NAFLD. This is in agreement with the diagnostic flow chart proposed by the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) Hepatology Committee (10). Most of these children were overweight or obese. All children were tested for secondary causes of hepatic steatosis, such as alcohol abuse, total parenteral nutrition, and chronic use of drugs known to induce hepatic steatosis (e.g., valproate, amiodarone and prednisone). Hepatitis A, B and C, cytomegalovirus, Epstein-Barr virus infections and coeliac disease were excluded according to appropriate serological tests.

Autoimmune liver disease, metabolic liver disease, Wilson's disease and alpha-1-antitrypsin-associated liver disease were ruled out using standard clinical, laboratory and histological criteria. Children with previously diagnosed type 1 diabetes were excluded from the study. All included children were Caucasians of Italian descent. None of them were taking any medication. The study protocol was approved by the ethics committee of the "Bambino Gesù" Children's Hospital. Written informed consent was obtained from the parents of each child.

Clinical and laboratory data

Body mass index (BMI) was measured as kilograms divided by the square of height in meters and then standardized BMI *z*-score was also calculated. Waist circumference was measured, with the patient in a standing position, on the horizontal plan between the lowest portion of the rib cage and the iliac crest. Blood pressure was measured on the right arm using a standard sphygmomanometer; the average of three blood pressure values was used for the analysis. Elevated blood pressure was defined by systolic or diastolic blood pressure >95th percentile for age and sex.

Venous blood samples were collected in the morning after an overnight fast of at least 8 hours. Measurements of serum glucose, creatinine (measured using a Jaffé rate blanked and compensated assay), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyltransferase (GGT) and other biochemical blood parameters were obtained using standard laboratory procedures at the central Laboratory of our hospital. Low-density lipoprotein (LDL)-cholesterol was calculated using the Friedewald's equation. Insulin levels were measured by a chemiluminescent immunoassay. Homeostasis model assessment (HOMA-IR) score (that includes both fasting glucose and insulin concentrations in its equation) was used for estimating insulin

resistance. Serum C-reactive protein (hs-CRP) was measured by a high-sensitivity latex agglutination method on HITACHI 911 Analyzer (Sentinel Ch., Milan). Serum tumor necrosis factor (TNF)-α and interleukin (IL)-6 levels were measured by sandwich enzyme-linked immunosorbent assays (R&D System Europe Ltd, Abingdon, UK). Glomerular filtration rate was estimated using the Bedside Schwartz equation [e-GFR= 0.413 x (height/Scr) where height is expressed in centimeters and serum creatinine in mg/dL], which has been shown to be accurate for estimating GFR in pediatric populations (11). A 24-hour urinary protein excretion was also measured by a commercially radioimmunoassay kit in a subgroup (n=109) of these children. This measurement was performed during hospitalization to minimize errors in the evaluation of urinary volume and sample retention. In accordance with widely used criteria, impaired fasting glycemia was defined by fasting glucose levels from 100 mg/dL to 125 mg/dL (5.6-6.9 mmol/L). The presence of type 2 diabetes was defined as a self-report of physician diagnosis, or a fasting glucose level ≥126 mg/dL (≥7.0 mmol/L).

Liver histology

Liver biopsies were performed using an automatic core biopsy 18-gauge needle (Biopince, Amedic, Sweden) under general anesthesia and ultrasound guidance, and then processed as described elsewhere (12). The main histological features commonly described for NAFLD, including hepatic steatosis, inflammation (portal and lobular), hepatocyte ballooning and fibrosis, were scored according to the scoring system for NAFLD developed by the National Institutes of Health–sponsored NASH Clinical Research Network (13). Briefly, hepatic steatosis was graded on a four-point scale: (0) steatosis involving fewer than 5% of hepatocytes, (1) steatosis involving up to 33% of hepatocytes, (2) steatosis involving 33% to 66% of hepatocytes, and (3) steatosis involving more than 66% of hepatocytes. Lobular inflammation was graded on a four-point scale: (0) no foci, (1)

fewer than two foci per 200× field, (2) two to four foci per 200× field, and (3) more than four foci per 200× field. Hepatocyte ballooning was graded from 0 to 2: (0) no balloon cells, (1) few balloon cells, and (2) many/prominent balloon cells. Portal tract inflammation was graded from 0 to 3 (0 is none, 1=mild, 2=moderate and 3=severe). The stage of hepatic fibrosis was quantified with a five-point scale: (0) no fibrosis, (1) perisinusoidal or periportal fibrosis [(1a) mild, zone 3, perisinusoidal; (1b) moderate, zone 3, perisinusoidal; and (1c) portal/periportal], (2) perisinusoidal and portal/periportal fibrosis, (3) bridging fibrosis, and (4) cirrhosis. Based on the NASH Clinical Research Network (13), we calculated the NAFLD Activity Score (NAS9, which is defined as the unweighted sum of the scores for steatosis (0-3), lobular inflammation (0-3), and ballooning (0-2); thus ranging from 0 to 8. Liver fibrosis is not included as a component of this activity score. Cases with NAS of 0 to 2 were largely considered not diagnostic of NASH; on the other hand, most cases with scores of ≥5 were diagnosed as definite NASH. Cases with activity scores of 3 and 4 were diagnosed as indeterminate or borderline NASH (13). In this study, we also used a newer grading histological score recently validated for pediatric NAFLD that has excellent performance for diagnosing NASH (14), i.e., the PNHS score which is calculated by using the weighted sum of hepatic steatosis (1-3), ballooning (0-2), lobular inflammation (0-3) and portal inflammation (0-2), according to the following formula: PNHS= 100 x exp (*z*PNHS)/[exp(*z*PNHS)], where *z*PNHS = -8.4 + 2.5 x steatosis + 3.5 x ballooning + 3.4 x lobular inflammation + 0.87 x portal inflammation.

PNPLA3 genotyping

Blood samples were collected from all children and the genomic DNA was extracted by QIAamp Blood MiniKit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The rate of success in extracting DNA was 100% for each study group. The *PNPLA3* rs738409 C>G SNP, encoding I148M, was genotyped with a 5'-nuclease TaqMan assay (Assay on Demand for

rs738409, Applied Biosystems, Foster City, CA) by personnel unaware of the clinical status of the patients. Post–polymerase chain reaction allelic discrimination was carried out through the measurement of allele-specific fluorescence on the Opticon 2 detection system (MJ Research, Waltham, MA). Random samples were confirmed by direct genotyping (Applied Biosystems 3500 Genetic Analyzer), which provided concordant results in all cases; controls were included in all analyzed batches, and quality controls were used to verify the reproducibility of the results (15). Distribution of the genotype was in Hardy-Weinberg equilibrium and the call rate exceeded 99%.

Statistical analysis

Due to the exploratory (hypothesis-generating) design of the study, it was not possible to compute a sample size. Data are expressed as means±SD, medians (inter-quartile ranges, IQR) or frequencies. The *PNPLA3* rs738409 associations were assessed using an additive genetic model. The chi-squared test or the Fisher's exact test (when appropriate) for categorical variables, the one-way ANOVA for normally distributed continuous variables and the Kruskal-Wallis test for non-normally distributed variables were used to test the differences in clinical and biochemical characteristics as well as in histologic features of NAFLD among children and adolescents, stratified by *PNPLA3* rs738409 polymorphism (**Tables 1 and 2**). The association between *PNPLA3* rs738409 variant and e-GFR values (included as a continuous measure, i.e. for each SD decrement) was tested using both an unadjusted linear regression model and the following three multivariable linear regression models: (*i*) model 1 adjusted for age and sex; (*ii*) model 2 adjusted for age, sex, waist circumference, systolic blood pressure and HOMA-IR score, and (*iii*) model 3 further adjusted for presence of NASH (i.e., defined as NAS ≥5) and histologic stage of liver fibrosis (**Table 3**). We did not additionally adjust for fasting glucose levels (or presence of impaired fasting glycemia), because this variable was already included in the formula of the HOMA-IR score. Using the *PNPLA3*

rs738409 variant as the exposure variable, the odds of having stage 2 CKD (i.e., defined as e-GFR value between 90 and 60 mL/min/1.73 m^2) among children with NAFLD were also determined using logistic regression analysis with the presence of stage 2 CKD as the binary outcome (**Table 4**). In this analysis, the presence of G/C and C/C *PNPLA3* rs738409 genotypes were combined into a single category, because the number of children with stage 2 CKD was relatively small (n=17). Finally, we also tested the independent association between *PNPLA3* rs738409 variant and 24-hour urinary protein excretion (included as a continuous measure, i.e. for each SD increment) by using a multivariable linear regression analysis that included the same list of the aforementioned covariates *plus* e-GFR (**Table 5**). Covariates included in all multivariable linear or logistic regression models were selected as potential confounding factors based on their significance in univariable regression analyses or based on their biological plausibility. A *p*-value <0.05 was considered statistically significant. Statistical analyses were performed using STATA software, version 14.2 (STATA, College Station, Texas, USA).

RESULTS

Among the 142 Caucasian children and adolescents with biopsy-proven NAFLD included in the study (74 boys and 68 girls; mean±SD: age 12±3 years; BMI 27.1±5 kg/m^2), 45 (31.7%) had G/G, 56 (39.4%) had G/C and 41 (28.9%) had C/C *PNPLA3* rs738409 genotype, respectively. These children and adolescents had a mean e-GFR value of 114.7±21 mL/min/1.73 m^2 (range: 70-180 mL/min/1.73 m^2) and a mean 24-hour proteinuria of 52.3±22 mg/day (range: 8-120 mg/day), respectively. Seventeen (11.9%) of these children had a stage 2 CKD with a mild decrease of the e-GFR value between 90 and 70 mL/min/1.73 m^2. Sixteen (11.3%) children had mildly elevated blood pressure values, and 10 (7%) of them also had impaired fasting glycemia (n=9) or type 2 diabetes

(n=1). All children included in the study did not take any drugs (including also drugs for hypertension, diabetes or dyslipidemia). According to the NAS score, 67 (47.2%) children had 'definite' NASH, whereas the remaining 75 (52.8%) subjects had 'not NASH'.

Table 1 shows the clinical and biochemical characteristics of children, stratified by *PNPLA3* rs738409 polymorphism. Compared to those with either G/C or C/C genotypes, children with G/G genotype were more likely to be older, centrally obese and more insulin resistant (as reflected by higher fasting insulin levels and greater HOMA-IR score) and had higher values of systolic blood pressure, serum triglycerides and creatinine. Sex, diastolic blood pressure, presence of impaired fasting glycemia/diabetes and circulating values of fasting glucose, hs-CRP, IL-6, TNF-alpha, liver enzymes and lipids (except for triglycerides) did not significantly differ among the three groups of children.

Notably, as shown both in **Table 1** and in **Figure 1**, the values of e-GFR (estimated by the Bedside Schwartz equation) decreased and 24-hour proteinuria increased progressively in relation to the *PNPLA3* rs738409 variant (p=0.002 and p=0.012 for trends by one-way ANOVA, respectively).

Table 2 shows the histologic features of NAFLD by *PNPLA3* rs738409 genotype. Among children with NAFLD, definite NASH was present in 17.1% of children with C/C genotype, 44.6% of those with G/C genotype, and 77.8% of those with G/G genotype (p<0.001). Children with G/G genotype also had higher degrees of hepatic steatosis, lobular inflammation, ballooning degeneration and fibrosis compared with those with either G/C genotype or C/C genotype. No significant difference was found in portal inflammation among the three groups of children.

As reported in **Table 3**, the *PNPLA3* G/G genotype was significantly associated with lower e-GFR, even after adjustment for age, sex, waist circumference, systolic blood pressure and HOMA-IR score (model 2). After further adjustment for presence of NASH and fibrosis stage (model 3), the results remained essentially unchanged, showing that the *PNPLA3* G/G genotype was closely associated with lower e-GFR. As also shown in the table, similar results were found for the *PNPLA3* G/C genotype. In these regression models, we did not additionally adjust for fasting glucose levels because this variable was already included in the HOMA-IR score. However, the results of the aforementioned model 3 did not change even when children with impaired fasting glycemia or diabetes were excluded from the analysis. Almost identical results were also found when we further adjusted for BMI *z* score and plasma triglyceride levels, or when we used the PNHS score, instead of the NAS score, for diagnosing NASH (data not shown).

Table 4 shows the association between the presence of *PNPLA3* rs738409 variant and risk of prevalent stage 2 CKD, and the effect of progressive adjustment(s) for established renal risk factors and potential confounding variables. In univariable logistic regression analysis (unadjusted model), the presence of *PNPLA3* C/G or G/G genotype was significantly associated with a ~four fold increased risk of having stage 2 CKD (unadjusted odds ratio 4.44, 95%CI 1.68-11.4). Notably, adjustment for age, sex, waist circumference, systolic blood pressure, HOMA-IR score and presence of NASH and histologic stage of liver fibrosis (model 3) did not substantially modify these results.

Table 5 shows the association between the risk allele (G) of *PNPLA3* rs738409 and 24-hour proteinuria (although this latter measurement was available only in 109 children). Similarly to the results above, the *PNPLA3* G/G genotype was associated with increasing 24-hour proteinuria, after

adjustment for age, sex, waist circumference, systolic blood pressure, HOMA-IR score and e-GFR (model 2), and even after further adjustment for the presence of NASH and histologic stage of liver fibrosis (model 3). These results remained essentially unchanged even when we excluded children with impaired fasting glycemia/diabetes from the analysis, or when we used the PNHS score, instead of the NAS score, for diagnosing NASH (data not shown).

DISCUSSION

To our knowledge, this is the first cross-sectional study aimed at examining the association between *PNPLA3* rs738409 polymorphism (i.e., the major genetic variant associated with susceptibility to NAFLD development and progression) and kidney function in a pediatric population with histologically confirmed NAFLD.

The novel findings of our study were that the presence of the risk allele (G) of rs738409 was significantly associated with decreasing e-GFR (as estimated by the Bedside Schwartz equation, which is currently considered the best method for estimating GFR in children aged 1-17 years [11]) and increasing 24-hour urinary protein excretion in a sample of Caucasian overweight children and adolescents with biopsy-proven NAFLD and normal or near-normal kidney function and without over proteinuria. Notably, these associations remained statistically significant even after adjusting for sex, age, measures of adiposity (waist circumference and BMI), systolic blood pressure, HOMA-IR score, and interestingly, these associations were also independent of the histologic severity of NAFLD.

The present findings confirm and expand the results of three small observational studies supporting the existence of a significant association between *PNPLA3* rs738409 polymorphism and decreased kidney function in middle-aged and elderly individuals with or without NAFLD (8,9,16). In a preliminary analysis of 202 Italian nondiabetic middle-aged individuals (about a third of whom had biopsy-confirmed NAFLD), Musso *et al.* showed that the G allele of rs738409 was significantly associated with lower e-$GFR_{CKD\text{-}EPI}$ and with a greater prevalence of CKD both in individuals with and in those without NAFLD (8). In another study involving 740 Japanese elderly individuals, Oniki *et al.* reported that non-obese carriers of the *PNPLA3* G/G genotype had lower e-$GFR_{CKD\text{-}EPI}$ compared with those carrying the C/C genotype, independently of common renal risk factors and coexisting ultrasound-detected NAFLD (9). Notably, in a subgroup of nearly 350 individuals who were followed for a mean period of 5.5 years, the authors also showed that the *PNPLA3* G/G genotype was associated with a significant e-GFR decline over the follow-up period (9). However, the generalizability of the findings of this study to other ethnic groups remains untested. Finally, we recently found that regardless of the coexistence of ultrasound-defined NAFLD and common renal risk factors, the *PNPLA3* G/G genotype was significantly associated with a higher prevalence of CKD in Italian post-menopausal women with established type 2 diabetes (16).

At first glance, one might question the clinical relevance of the small, although highly significant, differences in both e-GFR and 24-hour urinary protein excretion, across *PNPLA3* genotypes. However, it is important to stress that our findings are in a pediatric population, and as a function of their young age, children will be exposed over many years to multiple cumulative insults that will adversely affect their kidney function. Additionally, we have observed a significant and independent association between the presence of *PNPLA3* rs738409 variant and e-GFR values even when we categorized our children according to the severity of stage 2 CKD. That said, we

believe that the findings of our study (together with those derived from the three aforementioned studies in adults [8,9,16]), might have important clinical implications. Specifically, our findings support the assertion that *PNPLA3* genotyping may be useful not only for identifying individuals with NAFLD who are at increased risk of NASH (6,7,17,18), but *PNPLA3* genotyping may also be useful for identifying individuals with NAFLD who are at higher risk of kidney dysfunction. Thus, although these findings have to be further confirmed in larger cohorts, these results also suggest that *PNPLA3* genotyping could be useful to promote the implementation of specific prevention programs and treatment strategies for CKD among carriers of the *PNPLA3* rs738409 G/G genotype.

It is well established that the *PNPLA3* gene encodes a trans-membrane polypeptide chain exhibiting triglyceride hydrolase activity, which is mainly expressed on endoplasmic reticulum and lipid membranes of both hepatocytes and adipocytes (19-21). In particular, the *PNPLA3* rs738409 gene variant (i.e., encoding the isoleucine to methionine variant at protein position 148) impairs the phospholipase activity of the enzyme and accumulates at the surface of lipid droplets, where it acquires the ability to alter the phospholipid and triglyceride turnover(s), thus promoting liver fat accumulation (19-21). Currently, it is known that *PNPLA3* rs738409 polymorphism is associated with a greater predisposition to NASH and progressive liver fibrosis in pediatric and adult populations (6,7,15,22). In contrast, the possible biological mechanisms underlying the association between the G allele of rs738409 and decreasing kidney function are poorly understood. The most obvious explanation for our findings could be that the significant associations we observed between the G allele of rs738409 and decreasing e-GFR or raising 24-hour proteinuria are a consequence of both the shared renal/metabolic risk factors and the histological severity of NAFLD. However, it should be noted that in our study these associations remained statistically significant after adjustment for age, sex, systolic blood pressure, measures of adiposity and

HOMA-IR score. Importantly, these associations remained significant even after adjustment for the histological severity of NAFLD (i.e., NASH and fibrosis stage). Thus, it is plausible to speculate that the G allele of rs738409 might also have a direct adverse effect on the kidney. In support of this speculation, accumulating experimental evidence indicates that PNPLA3 is highly expressed in sinusoidal pericytes that store retinol and regulate its release in response to various metabolic signals (23), and furthermore, the *PNPLA3* rs738409 gene variant is associated with an increased release of multiple proinflammatory and profibrogenic factors (24). To date, renal pericytes have also become an intensively studied cell population in kidney biology and pathophysiology (25). Renal pericytes are stromal cells partially covering capillary walls that play an important role in angiogenesis and in regulating renal medullary and cortical blood flows, and may also serve as progenitors of interstitial myofibroblasts for the development of kidney fibrogenesis and glomerulosclerosis (25,26). Renal pericytes may interact with endothelial cells through distinct signaling pathways and their activation and detachment from capillaries after acute or chronic kidney injury is a critical event for driving CKD progression (25,26). Although it is currently unclear whether *PNPLA3* is directly implicated in the regulation of functions of renal pericytes, it is reasonable to hypothesize that the G allele of rs738409 might favor an unbalanced activation of pericytes in the kidney, thus promoting kidney fibrosis (27). Indeed, kidney fibrosis (i.e., the underlying pathological process of CKD) is a complex and progressive process, in which a myriad of acquired and genetic factors and molecular mediators (including also various phospholipases, such as PLA2 and PLD4) are involved (25,28). However, we acknowledge that further mechanistic studies are needed to better understand the role of the *PNPLA3* rs738409 polymorphism on the development of glomerular and interstitial fibrosis.

Our study has some important limitations that should be considered. First, its cross-sectional design does not allow for establishing the temporality and causality of the observed associations, although this limitation is mitigated by the fact that the genetic variant under study is inherited and, therefore, reverse causation does not apply. Second, it is possible that there is a selection bias from including Caucasian children/adolescents with biopsy-proven NAFLD (e.g., partly justifying the high prevalence of carriers with the *PNPLA3* G/G genotype we observed in our study). Hence, our results might not be necessarily generalizable to other pediatric and adult populations, although these findings (as discussed above) are also in line with those recently published in middle-aged and elderly individuals with or without NAFLD (8,9,16). Finally, although we did not directly measure GFR (by plasma iohexol disappearance), we used the most widely accepted prediction formula for estimating GFR in the pediatric population (i.e., the Bedside Schwartz formula) and not the CKD-epidemiology collaboration (EPI) formula, which is largely used for estimating kidney function in adults, because this latter formula has not been validated for the use in childhood and adolescence (29).

Notwithstanding these limitations, our study has some important strengths, including the relatively large sample size, the consecutive enrollment of the study population, the completeness of the database, the adjustment for common renal risk factors, and the use of liver biopsy for diagnosing and staging NAFLD disease severity.

In conclusion, this study shows for the first time that the G allele of rs738409 was strongly associated with both decreasing e-GFR and increasing 24-hour urinary protein excretion in Caucasian children and adolescents with biopsy-proven NAFLD. These associations were independent of age, sex, systolic blood pressure, measures of adiposity, HOMA-IR score and,

importantly, also the histological severity of NAFLD. However, more research is needed to better elucidate the link between *PNPLA3* rs738409 polymorphism and kidney dysfunction.

REFERENCES

1. Fleet SE, Lefkowitch JH, Lavine JE. Current concepts in pediatric nonalcoholic fatty liver disease. Gastroenterol Clin North Am 2017; 46: 217-231.
2. Mann JP, Valenti L, Scorletti E, Byrne CD, Nobili V. Nonalcoholic fatty liver disease in children. Semin Liver Dis 2018; 38: 1-13.
3. Selvakumar PKC, Kabbany MN, Nobili V, Alkhouri N. Nonalcoholic fatty liver disease in children: hepatic and extrahepatic complications. Pediatr Clin North Am 2017; 64: 659-675.
4. Mantovani A, Zaza G, Byrne CD, Lonardo A, Zoppini G, Bonora E, et al. Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: a systematic review and meta-analysis. Metabolism 2018; 79: 64-76.
5. Targher G, Byrne CD. Non-alcoholic fatty liver disease: an emerging driving force in chronic kidney disease. Nat Rev Nephrol 2017; 13: 297-310.
6. Anstee QM, Seth D, Day CP. Genetic factors that affect risk of alcoholic and nonalcoholic fatty liver disease. Gastroenterology 2016; 150: 1728-1744.
7. Macaluso FS, Maida M, Petta S. Genetic background in nonalcoholic fatty liver disease: A comprehensive review. World J Gastroenterol 2015; 21: 11088-11111.
8. Musso G, Cassader M, Gambino R. PNPLA3 rs738409 and TM6SF2 rs58542926 gene variants affect renal disease and function in nonalcoholic fatty liver disease. Hepatology 2015; 62: 658-659.

9. Oniki K, Saruwatari J, Izuka T, Kajiwara A, Morita K, Sakata M, et al. Influence of the PNPLA3 rs738409 polymorphism on non-alcoholic fatty liver disease and renal function among normal weight subjects. PLoS One 2015; 10: e0132640.
10. Vajro P, Lenta S, Socha P, Dhawan A, McKiernan P, Baumann U, et al. Diagnosis of nonalcoholic fatty liver disease in children and adolescents: position paper of the ESPGHAN Hepatology Committee. J Pediatr Gastroenterol Nutr 2012; 54: 700-713.
11. Schwartz GJ, Muñoz A, Schneider MF, Mak RH, Kaskel F, Warady BA, et al. New equations to estimate GFR in children with CKD. J Am Soc Nephrol 2009; 20: 629-637.
12. Manco M, Bedogni G, Marcellini M, Devito R, Ciampalini P, Sartorelli MR, et al. Waist circumference correlates with liver fibrosis in children with non-alcoholic steatohepatitis. Gut 2008; 57:1283-1287.
13. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al; Nonalcoholic Steatohepatitis Clinical Research Network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005; 41: 1313-1321.
14. Alkhouri N, De Vito R, Alisi A, Yerian L, Lopez R, Feldstein AE, Nobili V. Development and validation of a new histological score for pediatric non-alcoholic fatty liver disease. J Hepatol 2012; 57: 1312-1318.
15. Valenti L, Alisi A, Galmozzi E, Bartuli A, Del Menico B, 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.
16. Mantovani A, Zusi C, Sani E, Colecchia A, Lippi G, Zaza G, et al. Association between *PNPLA3* rs738409 polymorphism and decreased kidney function in postmenopausal type 2 diabetic women with or without non-alcoholic fatty liver disease. Diabet Metab 2019; *in press*.

17. Lonardo A, Bellentani S, Argo CK, Ballestri S, Byrne CD, Caldwell SH, et al. Epidemiological modifiers of non-alcoholic fatty liver disease: focus on high-risk groups. Dig Liver Dis 2015; 47: 997-1006.
18. European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO). EASL-EASD-EASO clinical practice guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016; 64: 1388-1402.
19. 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.
20. Donati B, Motta BM, Pingitore P, Meroni M, Pietrelli A, Alisi A, et al. The rs2294918 E434K variant modulates patatin-like phospholipase domain-containing 3 expression and liver damage. Hepatology 2016; 63: 787-798.
21. Mitsche MA, Hobbs HH, Cohen JC. Patatin-like phospholipase domain-containing protein 3 promotes transfer of essential fatty acids from triglycerides to phospholipids in hepatic lipid droplets. J Biol Chem 2018; 293: 6958-6968.
22. Seko Y, Yamaguchi K, Itoh Y. The genetic backgrounds in nonalcoholic fatty liver disease Clin J Gastroenterol 2018; 11: 97-102.
23. Pirazzi C, Valenti L, Motta BM, Pingitore P, Hedfalk K, Mancina RM, et al. PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells. Hum Mol Genet 2014; 23: 4077- 4085.
24. Bruschi FV, Claudel T, Tardelli M, Caligiuri A, Stulnig TM, Marra F, et al. The PNPLA3 I148M variant modulates the fibrogenic phenotype of human hepatic stellate cells. Hepatology 2017; 65: 1875-1890.

25. Kramann R, Humphreys BD. Kidney pericytes: roles in regeneration and fibrosis. Semin Nephrol 2014; 34: 374-383.
26. Shaw I, Rider S, Mullins J, Hughes J, Péault B. Pericytes in the renal vasculature: roles in health and disease. Nat Rev Nephrol 2018; 14: 521-534.
27. Romeo S, Dongiovanni P, Petta S, Pihalajamaki J, Valenti L. Reply: To PMID 25251399. Hepatology 2015; 62: 660.
28. Duffield JS. Cellular and molecular mechanisms in kidney fibrosis. J Clin Invest 2014; 124: 2299-2306.
29. Mian AN, Schwartz GJ. Measurement and estimation of glomerular filtration rate in children. Adv Chronic Kidney Dis 2017; 24: 348-356.

Table 1. Main clinical and biochemical characteristics of children and adolescents with biopsy-proven NAFLD, stratified by *PNPLA3* rs738409 genotype.

	CC genotype (*n*= 41)	**GC genotype (*n*= 56)**	**GG genotype (*n*= 45)**	***P* value**
Age (years)	10.7 ± 2.5	11.9 ± 3.5	13.1 ± 2.7	0.001
Male sex (%)	43.9	60.7	48.9	0.23
Body weight (kg)	50.9 ± 15	66.8 ± 25	73.3 ± 16	<0.001
BMI (kg/m^2)	23.5 ± 3.4	28.2 ± 5.9	29.3 ± 3.1	<0.001
Waist circumference (cm)	75.9 ± 14	87.3 ± 14	90.4 ± 10	<0.001
Systolic blood pressure (mmHg), n=118	107.5 ± 14	113.1 ± 13	117.3 ± 11	0.006
Diastolic blood pressure (mmHg), n=118	61.0 ± 11	62.6 ± 11	62.6 ± 10	0.78
IFG/Type 2 diabetes (%)*	0.0	3.6	4.4	0.42
Total cholesterol (mg/dL)	158.5 ± 36	156.8 ± 24	154.1 ± 29	0.79
LDL cholesterol (mg/dL)	101.9 ± 32	98.9 ± 26	96.9 ± 25	0.70
HDL cholesterol (mg/dL)	46.2 ± 10	46.4 ± 10	42.7 ± 10	0.15
Triglycerides (mg/dL)	83 (60-112)	80 (58-100)	104 (81-175)	0.001
Fasting glucose (mg/dL)	83.2 ± 9.0	85.4 ± 13.2	82.9 ± 8.5	0.44
Fasting insulin (mIU/L)	14.6 (10-19)	15.7 (12-23)	20.3 (14-30)	0.016
HOMA-IR score	3.0 (2.1-3.9)	3.4 (2.2-5.0)	4.2 (2.8-6.2)	0.003
AST (IU/L)	29 (23-33)	27 (23-33)	25 (23-32)	0.39
ALT (IU/L)	25 (19-40)	26 818-47)	28 (19-39)	0.51
GGT (IU/L)	14 (10-17	19 ± 13	20 ± 11	0.70
hs-CRP (mg/L)	1.3 (0.5-3)	0.5 (0.5-4)	0.7 (0.5-2)	0.14
IL-6 (pg/mL)	10.8 (7.9-18)	10.2 (8-14.7)	9.5 (8-12.3)	0.35
TNF-alpha (pg/mL)	6.9 (5.6-9)	6.7 (5.2-8.9)	7.2 (6.2-8.4)	0.21
Creatinine (mg/dL)	0.49 ± 0.10	0.56 ± 0.12	0.63 ± 0.13	<0.001
e-GFR$_{\text{Bedside Schwartz}}$ (mL/min/1.73 m^2)	125.3 ± 23	112.8 ± 18	107.5 ± 20	0.002
Proteinuria (mg/24h), n=109	42.9 ± 20	53.9 ± 22	58.5 ± 21	0.012

Sample size, *n*=142 unless where indicated. Data are expressed as means ± SD, medians (IQR) or relative frequencies. Differences among the three groups of children were tested by the chi-squared test or the Fisher's exact test for categorical variables, the one-way ANOVA for normally distributed continuous variables, and the Kruskal-Wallis test for non-normally distributed continuous variables (i.e., fasting insulin, HOMA-IR score, triglycerides, C-reactive protein, IL-6, TNF-alpha, AST and ALT levels). *Impaired fasting glycemia (IFG) was present in 9 children and was defined as a fasting glucose level between 100 mg/dL and 125 mg/dL; only a subject belonging to the GC genotype group had established type 2 diabetes.

Abbreviations: ALT, alanine aminotransferase, AST, aspartate aminotransferase; BMI, body mass index; hs-CRP, high sensitivity C reactive protein; e-GFR, estimated glomerular filtration rate; GGT, gamma-glutamyltransferase; HOMA-IR, homeostasis model assessment-insulin resistance; IFG, impaired fasting glycemia; IL-6, interleukin-6; TNF, tumor necrosis factor.

Table 2. Liver histology characteristics of children and adolescents with NAFLD, stratified by *PNPLA3* rs738409 genotype.

	CC genotype (*n*= 41)	**GC genotype (*n*= 56)**	**GG genotype (*n*= 45)**	***P* value**
Steatosis grade				<0.005
5%-33%	16 (39%)	16 (28.6%)	10 (22.2%)	
33% - 66%	21 (51.2%)	27 (48.2%)	15 (33.4%)	
>66%	4 (9.8%)	13 (23.2%)	20 (44.4%)	
Lobular inflammation				<0.005
No foci	13 (31.7%)	9 (16.1%)	2 (4.4%)	
<2 foci per 200 x field	28 (68.3%)	42 (75.0%)	36 (80.0%)	
2-4 foci per 200 x field	0 (0%)	5 (8.9%)	7 (15.6%)	
Portal inflammation				0.96
No	3 (7.3%)	4 (7.1%)	3 (6.7%)	
Mild	33 (80.5%)	47 (83.9%)	38 (84.4%)	
Moderate	5 (12.2%)	5 (8.9%)	4 (8.9%)	
Ballooning degeneration				<0.001
None	14 (34.1%)	15 (26.8%)	3 (6.7%)	
Few	22 (53.7%)	21 (37.5%)	17 (37.8%)	
Many	5 (12.2%)	20 (35.7%)	25 (55.5%)	
Fibrosis stage				<0.01
No fibrosis	8 (19.5%)	10 (17.9%)	1 (2.2%)	
Perisinusoidal fibrosis	27 (65.8%)	40 (71.4%)	30 (66.7%)	
Periportal fibrosis	6 (14.6%)	6 (10.7%)	14 (31.1%)	
Bridging fibrosis	0 (0%)	0 (0%)	0 (0%)	
Definite NASH*	7 (17.1%)	25 (44.6%)	35 (77.8%)	<0.001

Simple size, n=142. Data are expressed as absolute and relative frequencies. Differences among the three groups of children were tested by the Fisher's exact test. *Definite NASH was classified as the presence of a NAFLD Activity Score ≥5 according to the NASH Clinical Research Network classification.

Table 3. Association between *PNPLA3* rs738409 genotype and estimated glomerular filtration rate in children and adolescents with biopsy-proven NAFLD.

	***B* coefficient(s)**	**95% Confidence Interval**	***P* value**
Unadjusted Model			
e-GFR for 1-SD decrement (*i.e.,* 21 ml/min/1.73 m^2)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	-12.5	-20.6 to -4.31	0.003
GG genotype	-17.8	-26.3 to -9.23	<0.001
Adjusted Model 1			
e-GFR for 1-SD decrement (*i.e.,* 21 ml/min/1.73 m^2)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	-11.3	-19.7 to -3.01	0.008
GG genotype	-16.5	-25.5 to -7.51	<0.001
Adjusted Model 2 (n=118)			
e-GFR for 1-SD decrement (*i.e.,* 21 ml/min/1.73 m^2)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	-15.2	-25.4 to -4.99	0.004
GG genotype	-21.9	-33.2 to -10.7	<0.001
Adjusted Model 3 (n=118)			
e-GFR for 1-SD decrement (*i.e.,* 21 ml/min/1.73 m^2)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	-16.0	-26.5 to -5.55	0.003
GG genotype	-23.6	-36.3 to -10.8	<0.001

Sample size, *n*=142 unless where indicated. Data are expressed as beta coefficients and 95% confidence intervals as tested by multivariable linear regression analysis. Continuous values of e-GFR$_{\text{Bedside Shwartz}}$ (expressed as 1-SD decrement) were included as the dependent variable in multivariable linear regression analysis. *Ref.*= reference category.

Multivariable linear regression models were adjusted for the following covariates: model 1: age and sex; model 2: age, sex, waist circumference, systolic blood pressure and HOMA-IR score; and model 3: adjusted for the same covariates of model 2 *plus* presence of NASH (i.e., defined as a NAFLD Activity Score ≥5) and histologic stage of liver fibrosis.

Table 4. Association between *PNPLA3* rs738409 genotype and presence of stage 2 chronic kidney disease in children and adolescents with biopsy-proven NAFLD.

	Odds Ratio(s)	95% Confidence Interval	*P* value
Unadjusted Model			
e-GFR ≤90 ml/min/1.73 m^2			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC or GG genotype	4.44	1.68 - 11.4	0.003
Adjusted Model 1			
e-GFR ≤90 ml/min/1.73 m^2			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC or GG genotype	4.28	1.55 - 11.8	0.005
Adjusted Model 2 (n=118)			
e-GFR ≤90 ml/min/1.73 m^2			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC or GG genotype	5.39	1.63 - 17.8	0.005
Adjusted Model 3 (n=118)			
e-GFR ≤90 ml/min/1.73 m^2			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC or GG genotype	7.87	1.89 - 32.8	0.005

Sample size, *n*=142 unless where indicated. Data are expressed as odds ratio and 95% confidence intervals as tested by logistic regression analysis. Presence of stage 2 CKD (defined as an e-GFR$_{\text{Bedside Shwartz}}$ value of 90-60 mL/min/1.73 m^2; n=17 children) was included as the dependent variable in logistic regression analysis. *Ref*.= reference category.

Multivariable logistic regression models were adjusted for the following covariates: model 1: age and sex; model 2: age, sex, waist circumference, systolic blood pressure and HOMA-IR score; model 3: adjusted for the same covariates of model 2 *plus* presence of NASH (i.e., defined as a NAFLD Activity Score ≥5) and histologic stage of liver fibrosis.

Table 5. Association between *PNPLA3* rs738409 genotype and 24-hour urinary protein excretion in children and adolescents with biopsy-proven NAFLD.

	***B* coefficient(s)**	**95% Confidence Interval**	***P* value**
Unadjusted Model			
24-h Proteinuria for 1-SD increment (*i.e.*, 22 mg/24h)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	10.9	0.95 to 20.9	0.032
GG genotype	15.6	5.12 to 26.0	0.004
Adjusted Model 1			
24-h Proteinuria for 1-SD increment (*i.e.*, 22 mg/24h)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	8.89	-1.28 to +19.0	0.086
GG genotype	12.5	1.59 to 23.4	0.025
Adjusted Model 2 (n=91)			
24-h Proteinuria for 1-SD increment (*i.e.*, 22 mg/24h)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	12.8	0.53 to 25.1	0.041
GG genotype	20.1	6.58 to 33.7	0.004
Adjusted Model 3 (n=91)			
24-h Proteinuria for 1-SD increment (*i.e.*, 22 mg/24h)			
PNPLA3 rs738409			
CC genotype	*Ref.*	*Ref.*	*Ref.*
GC genotype	11.1	-1.39 to +23.7	0.080
GG genotype	15.3	1.12 to 30.5	0.046

Sample size, *n*=109 unless where indicated. Data are expressed as beta coefficients and 95% confidence intervals as tested by multivariable linear regression analysis. Continuous values of 24-hour proteinuria (expressed as 1-SD increment) were included as the dependent variable in multivariable linear regression analysis. *Ref*.= reference category.

Multivariable linear regression models were adjusted for the following covariates: model 1: age and sex; model 2: age, sex, waist circumference, systolic blood pressure, e-$GFR_{Bedside\ Schwartz}$ and HOMA-IR score; and model 3: adjusted for the same covariates of model 2 *plus* presence of presence of NASH (i.e., defined as a NAFLD Activity Score ≥5) and histologic stage of liver fibrosis.

FIGURE LEGENDS

Figure 1. Box plot of estimated glomerular filtration rate values (i.e., e-GFR as estimated by the Bedside Schwarz equation) of 142 children and adolescents with biopsy-proven NAFLD, stratified by *PNPLA3* rs738409 genotype. *P*-values for the trend by one-way ANOVA. A post-hoc Tukey's HSD (Honestly Significant Difference) test was used for pairwise comparisons of means among the three groups of children. The central rectangle spans the 1st quartile to the 3rd quartile (the interquartile range [IQR]). The segment inside the rectangle shows the median and "whiskers" above and below the box show the locations of 1.5 x IQR values.

Figure 2. Box plot of 24-hour proteinuria values of children and adolescents with biopsy-proven NAFLD, stratified by *PNPLA3* rs738409 genotype. *P*-values for the trend by one-way ANOVA. A post-hoc Tukey's HSD (Honestly Significant Difference) test was used for pairwise comparisons of means among the three groups of children. The central rectangle spans the 1st quartile to the 3rd quartile (the interquartile range [IQR]). The segment inside the rectangle shows the median and "whiskers" above and below the box show the locations of 1.5 x IQR values. Data of 24-hour proteinuria were available in 109 children.