HHS Public Access

Author manuscript

Acta Diabetol. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

Acta Diabetol. 2015 April ; 52(2): 373–382. doi:10.1007/s00592-014-0654-3.

Genome-wide analysis of hepatic lipid content in extreme obesity

Johanna K. DiStefano[1,*], **Christopher Kingsley**[1], **G. Craig Wood**[2], **Xin Chu**[2], **George Argyropoulos**[2], **Christopher D. Still**[2], **Stefania Cotta Doné**[1], **Christophe Legendre**[1], **Waibhav Tembe**[1], and **Glenn S. Gerhard**[2,3]

[1]Diabetes, Cardiovascular and Metabolic Diseases Division, Translational Genomics Research Institute, 445 Fifth Street, Phoenix, AZ 85004

[2]Geisinger Obesity Institute, Geisinger Clinic, 100 N. Academy Ave., Danville, PA 17822

[3]Department of Biochemistry and Molecular Biology, Institute for Personalized Medicine, Department of Pathology and Laboratory Medicine, Pennsylvania State University College of Medicine, Room C5750, 500 University Drive, MC - H171, Hershey, PA 17033

Abstract

Individuals with type 2 diabetes have an increased risk of developing nonalcoholic fatty liver disease (NAFLD), and NAFLD patients are also at greater risk for developing type 2 diabetes. Although the relationship between type 2 diabetes and NAFLD is highly interconnected, the pathogenic mechanisms linking the two diseases are poorly understood. The goal of this study was to identify genetic determinants of hepatic lipid accumulation through association analysis using histological phenotypes in obese individuals. Using the Illumina HumanOmniExpress BeadChip assay, we genotyped 2300 individuals on whom liver biopsy data were available. We analyzed total bilirubin levels, which are linked to fatty liver in severe obesity, and observed the strongest evidence for association with rs4148325 in *UGT1A* ($P<5.0 \times 10^{-93}$), replicating previous findings. We assessed hepatic fat level and found strong evidence for association with rs4823173, rs2896019, and rs2281135, all located in *PNPLA3* and rs10401969 in *SUGP1*. Analysis of liver

[*]Corresponding author: Please send all correspondence to: Johanna K. DiStefano, Ph.D., Translational Genomics Research Institute, 445 North Fifth Street, Phoenix, AZ 85004, Tel: 602.343.8812, FAX: 602.343.8844, jdistefano@tgen.org. Copy on all correspondence: Glenn S. Gerhard, M.D., Department of Biochemistry and Molecular Biology, Institute for Personalized Medicine, Dept of Pathology and Laboratory Medicine, Pennsylvania State University College of Medicine, Room C5750, 500 University Drive, MC - H171, Hershey, PA 17033, Tel: 717.531.3615, FAX: 717.531.7072, ggerhard@hmc.psu.edu.

Contribution statement

JKD developed study design, analyzed data, wrote manuscript, and provided final approval of the version to be published. CK analyzed genotype data. GCW performed data analysis. XC prepared samples. GA contributed to study design and reviewed manuscript. CDS contributed to study design and reviewed manuscript. QJ contributed to data analysis. SCD performed genotyping. CL and WT analyzed RNA-Seq data. GSG developed study design, analyzed data, and wrote manuscript.

Statement of Informed Consent

Informed consent was obtained from all patients for being included in the study.

Conflict of Interest:

Johanna K. DiStefano, Christopher Kingsley, G. Craig Wood, Xin Chu, George Argyropoulos, Christopher D. Still, Stefania Cotta Doné, Christophe Legendre, Waibhav Tembe, and Glenn S. Gerhard declare that they have no conflict of interest.

Statement of Human and Animal Rights

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5).

transcript levels of 20 genes residing at the *SUGP1/NCAN* locus identified a 1.6-fold change in expression of the *LPAR2* gene in fatty liver. We also observed suggestive evidence for association between low-grade fat accumulation and rs10859525 and rs1294908, located upstream from *SOCS2* and *RAMP3*, respectively. *SOCS2* was differentially expressed between fatty and normal liver. These results replicate findings for several hepatic phenotypes in the setting of extreme obesity and implicate new loci that may play a role in the pathophysiology of hepatic lipid accumulation.

Keywords

Obesity; hepatic lipid; GWAS; bilirubin; LPAR2; PNPLA3; SOCS2; RNA Sequencing; gastric bypass surgery

Introduction

Type 2 diabetes and obesity are associated with fat accumulation in the liver, which is commonly diagnosed as non-alcoholic fatty liver disease (NAFLD) and is associated with the metabolic syndrome [1]. Indeed, the majority of patients with NAFLD are overweight or obese, and individuals with type 2 diabetes have an increased risk of developing the disease [2]. Conversely, patients with fatty liver are also at greater risk for developing type 2 diabetes [2]. Although the relationship between type 2 diabetes and NAFLD is highly interconnected, the pathogenic mechanisms linking the two diseases are poorly understood. Similarly, hepatic lipid accumulation has been proposed as a cause, rather than a consequence, for the development of hepatic insulin resistance [3], but the molecular mechanism by which this occurs is not yet known.

Both environmental and genetic factors play a significant role in mediating risk for development of increased fat in the liver [4]. A major risk factor is obesity, and extreme obesity in particular, is associated with a high prevalence of hepatic fat accumulation [5]. A genetic component is supported by observations of familial aggregation, family history of type 2 diabetes, and ethnic differences in disease prevalence [6–9]. In addition, hepatic fat accumulation and liver fat fraction are also highly heritable [10]. A number of genetic variants have been associated with development of fatty liver [11], but the strongest evidence for association across multiple studies has been observed with marker rs738409 in the patatin-like phospholipase domain containing 3 gene (*PNPLA3*). The first findings of association were in participants of the Dallas Heart Study, in whom the G allele, which causes an isoleucine to methionine substitution at position 148 of the protein sequence, was associated with hepatic fat content estimated via proton magnetic resonance spectroscopy [12]. The association between the G allele and hepatic fat has been replicated across individuals from different ethnicities and geographical regions; however, the effects of the *PNPLA3* variant appear independent of features of the metabolic syndrome, including obesity, insulin resistance, and hypertriglyceridemia [12–19]. These results suggest that *PNPLA3*-associated increases in hepatic fat content may have different metabolic consequences than other mechanisms of lipid accumulation.

Most genetic studies of fatty liver have used image-based phenotyping rather than histological determination, which remains the gold standard for diagnosis and staging of the disease. Here, we investigated the genetic determinants of hepatic lipid content based upon biopsy-proven phenotype using a genome-wide approach. We utilized a population design of extreme obesity and a high level of adiposity to homogenize the potential for a fatty liver phenotype resulting from caloric overconsumption. That is, all patients will have consumed sufficient calories to induce liver fat accumulation in susceptible individuals. We first conducted a genome-wide analysis of serum total bilirubin level, strongly influenced by a single major locus (20,21), to determine the genetic architecture in extreme obesity. Bilirubin has also been associated with fatty liver in morbid obesity [20, 21], correlated with insulin sensitivity in animal models [22], and inversely related to insulin resistance in children and adolescents, and has been reported to be a risk factor for mortality in type 2 diabetes [23]. We then extended this approach to histologically determined hepatic lipid content as a phenotype. To minimize the effect of genetic variants that predispose to severe fat accumulation, we also conducted genome-wide analysis of mild or incipient hepatic lipid accumulation, which may have more relevance to the onset of hepatic insulin resistance. Finally, we examined levels of expression of two genes harboring variants associated with hepatic fat content in this extremely obese population, providing the first evidence for these loci in the pathophysiology of hepatic fat accumulation.

Methods

Study sample

Study participants were comprised of Caucasian individuals enrolled in the Bariatric Surgery Program at the Geisinger Clinic Center for Nutrition and Weight Management. Blood samples and medically related data were obtained from patients during a standardized, multi-disciplinary, pre-operative program. The preoperative preparation process included a separate clinic visit in which a 1–2 hour comprehensive clinical interview and psychological evaluation was conducted by certified psychology professionals that included assessment of drugs, alcohol, and smoking behaviors. Patients with definite or possible evidence of addictive behaviors were not allowed to proceed to surgery. In addition, any clinical data, e.g., diagnostic ICD-9 codes, medical and medication history, that indicated drug or alcohol abuse were identified and those patients removed from the study. In all participants, wedge biopsies of the liver were obtained intraoperatively from a similar anatomic location, approximately 10 cm lateral and to the left of the falciform ligament. The tissue was sectioned so that one-fourth to one-third was submerged directly in RNA*later* (Applied Biosystems/Ambion; Austin, TX), and the remainder was fixed in neutral buffered formalin, stained with hematoxylin and eosin, and histologically evaluated as part of clinical standard of care using NASH CRN criteria [24]. A second pathologist reviewed histological data. Hepatic lipid content was graded using low- to medium power evaluation of hepatocyte involvement by macrosteatosis and/or microsteatosis with grade 0 involving <5% of parenchyma, grade 1 involving 5%–33%, grade 2 involving 33%–66%, and grade 3 involving >66%.

Clinical variables were obtained from an electronic database as described previously [25], including basic clinical measures, demographics, diagnostic ICD-9 codes, medical and medication history, and common lab results. Research was approved by the Institutional Review Boards of the Geisinger Clinic and the Translational Genomics Research Institute, and informed consent was obtained from all patients for being included in the study.

Genome-wide SNP genotyping and quality control

Genomic DNA was isolated from EDTA anti-coagulated blood samples and genotyped using the Infinium HD Ultra BeadChip assay (Illumina; San Diego, CA). Over half of the more than 700,000 SNPs on the IlluminaOmniExpress BeadChip reside within 10 Kb of a RefSeq gene, while only about 15,000 are predicted to cause a non-synonymous amino acid change. Chips were imaged using the iScan system and results were analyzed with the GenomeStudio v. 2010.3 software program. Multi-dimensional scaling (MDS) was performed using the plink application (http://pngu.mgh.harvard.edu/purcell/plink/). Samples with MDS C1 >0.011 were considered outliers and removed from analysis. Samples with >10% missing genotype rates were also removed from analysis. SNPs were selected to have minor allele frequency >1% and a Hardy-Weinberg p-value $>10^{-6}$. A total of 605,718 SNPs were included for association analyses.

Statistical Analysis

We utilized a two-pronged strategy for association testing. First, we used linear regression to assess additive allelic effects. Linear regression was used to fit encoded genotypes to numerically graded level of hepatic fat or mean serum bilirubin level. We also used logistic regression to evaluate the association of genotype and grade 1 versus grade 0 hepatic lipid content. The unadjusted regression, false discovery rate (FDR), and Bonferroni-adjusted p-values were used to determine statistical significance. The eigenstrat method was used to analyze and adjust for any potential confounding based on unexpected differences in ancestry. Full-scan permutation was used with principal components analysis with 10 principal components calculated and full permutation testing performed five times to identify any outliers >6 standard deviations. Outliers were removed. All statistical analyses were performed using Plink version 1.07 and HelixTree SNP Variation Suite 7 (http://www.goldenhelix.com/). SNAP (SNP Annotation and Proxy Search) was used to calculate linkage disequilibrium [26].

RNA isolation, sample preparation, and sequencing

Total RNA was extracted from liver samples using the RNeasy kit (Qiagen; Valencia, CA). RNA sequencing was performed using the Illumina HiSeq2000 next generation sequencing platform. Output was divided across 24 individual flowcell lanes. We sequenced 72 samples with approximately 6 Gb (~60 million 108 bp paired-end reads) of sequencing yield dedicated to each sample. During base-calling, low quality reads were identified and removed, and indexed reads were identified and grouped accordingly. Quality scores per base were calculated and visualized. Filtered reads were then aligned with the entire human genome using the Bowtie program. Aligned RNA-Seq reads were imported into the Cufflinks program, which assembles alignments into a parsimonious set of transcripts and estimates relative abundance based on the number of reads per fragment, and we used the

TopHat program to provide base-pair resolution exon annotations along with approximate quantitation of expression for those exons.

Results

Genotyping was performed on DNA obtained from 2300 patients. Following filtering and removal of unsatisfactory data for call rate, minor allele frequency, Hardy-Weinberg equilibrium, data from a total of 2084 samples were available. Individuals with any evidence of viral hepatitis, HIV, or cancer (9 samples) and outliers following Eigenstrat principal component permutation testing (207 samples) were then removed. A summary of demographic and clinical characteristics of the final cohort (n=1868) is shown in Table 1. As is typical for patients enrolled in bariatric surgery programs [27], the population was primarily female (>81%). In addition, the percentage of males with grade 2 and 3 hepatic lipid content was significantly higher than in those without liver fat. The average age of 45 years, BMI of 47 kg/m^2, and waist circumference of 53 inches were not significantly different across grades. Mean values of glucose, insulin, and HbA1c levels were significantly higher in patients with any evidence of hepatic lipid accumulation (grades 1, 2, or 3). Mean triglyceride levels also increased with hepatic fat grade from 138 to >200 mg/dL in grades 2 and 3, as did mean ALT levels from 23 to 46 U/L, and mean AST from 23 to 36 U/L. However, total cholesterol, LDL-C, HDL-C, alkaline phosphatase, and total bilirubin did not change significantly with level of hepatic lipid accumulation.

We first investigated association between genetic variants and total bilirubin level. Previous studies have shown highly significant associations with variants in the family of uridine 5′-diphospho-glucuronosyltransferase genes, i.e., the *UGT1A* locus [23]. We performed this analysis to serve as a “replication control” for our data and, given the association between total bilirubin and aspects of type 2 diabetes, to conduct an analysis in the new context of extreme obesity. We observed a cluster of variants on chromosome 2 showing strong association with total bilirubin. The strongest evidence for association was observed with rs4148325 (6–10% decrease in mean bilirubin level per copy of T allele; nominal P=5.0 × 10^{-93}), located in the *UGT1A1* gene (Table 2). Interestingly, despite the relative homogeneity of the study sample, principal components analyses of genotype data excluded nearly 10% of the population, implying that clinical grade assessment of ethnic status does not preclude significant genetic population stratification. Potential stratification in other populations of apparent European ancestry may have similar implications for previously reported GWAS studies.

We next examined the relationship between genotype and level of hepatic fat based on histological hepatic lipid grade. As shown in Figure 1, four markers approached FDR- or Bonferroni-corrected statistical significance, rs4823173, rs2896019, and rs2281135, all in the *PNPLA3* gene, and rs10401969 which resides in intron 8 of the SURP and G patch domain containing 1 (*SUGP1*) gene located within the *NCAN* locus (Table 2). rs4823173 and rs2281135 are 3840 bp apart and show moderate linkage disquilibrium (R^2= 0.843). rs4823173 and rs2896019 are 4964 bp apart with identical linkage, whereas rs2896019 and rs2281135 are 1124 bp apart and in stronger linkage (R^2= 1.000). Given the known dominant nature of the *PNPLA3* locus on hepatic fat level, we reanalyzed the association

using a dominant model and observed statistically significant evidence for association between *PNPLA3* variants and hepatic fat grade (Table 2). For rs4823173, a dominant effect was evident, with different (P<0.0001) mean grades of hepatic fat of 1.08 (+/− 0.89) for the 1272 major allele (GG) homozygotes, and 1.32 (+/− 0.95) and 1.33 (+/− 0.86), respectively for the 557 GA and 52 AA minor allele carriers and homozygotes. In contrast, the effect on mean (+/1 SD) grade for rs10401969 varied (P<0.0001) from 1.11 (+/− 0.90), 1.41 (+/− 0.97), and 1.83 (+/− 1.17) for 0 (n= 1625), 1 (n= 251), and 2 (n= 6) copies of the minor C allele, respectively, consistent with an additive effect.

Because over 80% of the population was female, we also conducted sex-specific analyses. Genome-wide association analysis for 1512 female individuals using an additive model yielded similar but less significant results (Supplemental Figure S1), with the peak p-values for the *NCAN* and *SUGP1/PNPLA3* loci just less than 10^{-6}. A further diminution in statistical significance was present for the 356 males (Supplemental Figure S2). Interestingly, the peak P value (greater than ~10^{-4}) for the *PNPLA3* locus was disproportionately reduced relative to the *SUGP1/PNPLA3* locus, suggesting that the result was not due solely to reduced power, but that the locus may have a greater influence on hepatic fat accumulation in women compared to men.

The association of rs10401969 with hepatic lipid grade is consistent with previously reported findings of association in the *NCAN* locus [28]. In this locus, a number of variants spanning a genomic region encompassing approximately twenty genes have been shown to be nominally associated with hepatic fat. To date, the gene that contributes to the hepatic fat phenotype has not been identified. To determine whether any of the genes in this locus showed changes in expression with respect to hepatic lipid content, we sequenced RNA from biopsied liver tissue of 12 patients with normal histology and 12 with liver fat (steatosis grades 2 and 3). We evaluated the ratio of transcript numbers for *MEF2BNB-MEF2B, RFXANK, LPAR2, GMIP, SUGP1, ZNF14, PBX4, ATP13A1, MAU2, NCAN, CILP2, GATAD2A, TSSK6, ZNF101, NR2C2AP, NDUFA13, HAPLN4, TM6SF2, MEF2B, and YJEFN3* (Supplemental Table 1). Of these genes, *LPAR2* was the only gene whose transcript level was differentially expressed, with a 1.6-fold increase in fatty vs. normal liver (P=0.00015).

We next sought to investigate the genetic architecture of low grade hepatic lipid accumulation because 1) this condition affected the highest percentage of individuals in our cohort, 2) *PNPLA3* and *SUGP1*/*NCAN* locus variants that are strongly associated with moderate to severe grades of hepatic fat may potentially obscure other loci, and 3) genes associated with incipient hepatic lipid accumulation may be more likely to play a role in the development of insulin resistance. Few studies have conducted such subgroup analysis. We therefore performed logistic regression analysis using an additive genetic model with grade 1 and grade 0 as case and control groups, respectively (Fig. 2). In this analysis, we observed highly suggestive evidence for association (Table 2) with rs10859525 ($P=3.8 \times 10^{-7}$) on chromosome 12, located approximately 15 kb upstream from the suppressor of cytokine signaling 2 (*SOCS2*) gene and with rs1294908 ($P=7.1 \times 10^{-7}$) located on chromosome 7, residing ~10 kb downstream from the receptor (G protein-coupled) activity modifying protein 3 (*RAMP3*) gene. For SOCS2 SNP rs10859525, the mean hepatic fat grade was 0.75

(+/− 0.43) for the 454 major allele homozygotes (GG), 0.64 (+/− 0.48) for the 622 heterozygotes (GA), and 0.55 (+/− 0.50) for the 214 minor allele homozygotes (AA). The pattern for RAMP3 SNP rs1294908 genotypes was similar to that of SOCS2 with mean hepatic fat grade of 0.67 (+/− 0.48) for the 808 major allele homozygotes (AA), 0.56 (+/− 0.50) for the 426 heterozygotes, and 0.42 (+/− 0.50) for the 55 minor allele homozygotes (GG). *SOCS2* transcript levels were statistically decreased in RNA sequencing, with quantitative PCR (data not shown) from fatty liver samples (grades 1–3) showing about 30% of the transcript levels of normal liver (P=0.0002). In contrast, there were no significant differences in *RAMP3* transcript levels between the two groups.

Discussion

We sought to define the genetic architecture of histologically documented hepatic lipid accumulation in a population with extreme obesity. Our analysis focused on fatty liver because of the potential role of hepatic lipids in the development of insulin resistance and to provide more in-depth analysis of an understudied phenotype relative to more advanced stages of NAFLD/NASH. In addition, most genetic studies of fatty liver have used image-based phenotyping rather than histological determination. We analyzed bariatric surgery patients who undergo liver biopsy not because of a specific indication of potential liver disease, but as a standard of care due to the relatively safe opportunity to diagnose potentially severe occult liver disease resulting from fatty liver. Thus, the availability of histological data from a population in which liver biopsies would otherwise not ethically be performed provides a unique opportunity to explore biological mechanisms not possible with other study designs. This relatively unbiased biopsy data also yields information to further evaluate liver status and perform work-up for other potential liver diseases such as viral hepatitis or alcohol abuse, which can then be excluded from analysis. In addition, the extreme level of obesity also likely homogenizes the potential for a fatty liver phenotype from caloric over-consumption, i.e., essentially all patients have consumed sufficient calories to induce fatty liver in susceptible individuals. The extensive pre-operative psychological and medical preparation program for bariatric surgery includes in-depth evaluation of excess alcohol consumption and drug abuse, either of which exclude patients from consideration for bariatric surgery, thus minimizing two potential environmental factors that could influence the occurrence of liver fat.

Our results extend to a population with extreme obesity several known liver-related genetic loci. We observed strong evidence for association of total bilirubin with the *UGT1A1* locus that was first identified in Caucasians [20, 21], and later confirmed in in Chinese [29] and African-American populations [30]; however, our study is the first performed in the setting of extreme obesity. We also observed that total bilirubin level was not related to hepatic lipid grade, demonstrating that neither the degree of obesity nor the severity of fatty liver impacted the influence of the *UGT1A1* genetic variants. In contrast to these findings, we did not detect association between total bilirubin levels and *SLCO1B1* variants, which have also been associated with this phenotype [31]. However, the strength of the *SLCO1B1* association was much weaker relative to *UGT1A1* variants, and the current study may not have been sufficiently powered to detect such an association. Alternatively, this locus may

not play a role in the regulation of bilirubin levels in individuals with extreme levels of obesity.

We also replicated two loci associated with both CT- and biopsy-proven NAFLD in a genome wide association study (GWAS) of a Caucasian population: the *PNPLA3* and the *NCAN* loci [28]. *PNPLA3* variants have also been associated with NAFLD through GWAS in a Japanese population [32], but not in the NASH Clinical Research Network cohort [33], although analysis of *PNPLA3* as a candidate gene using a large unphenotyped control population was demonstrated [34]. The causal variant in *PNPLA3*, rs738409, which encodes an I148M substitution, has a strong, dominant effect on liver fat, even under the condition of extreme obesity [35, 36]. Unfortunately, rs738409 was not present on the genotyping chip we used for this study. It is, however, in moderate to strong linkage disequilibrium with other *PNPLA3* variants (rs4823173, rs2896019, and rs2281135) that were genotyped and these markers reside together in the same haplotype block (Supplemental Figure 1).

We also observed association between hepatic grade and variants in the *SUGP1/NCAN* locus. While attempts have been made to identify the causative gene in this locus that underlies the association with fatty liver [37], no candidate has yet been confirmed. Recently, a variant (rs58542926) coding for a missense mutation (Glu167Lys) in the transmembrane 6 superfamily member 2 (*TM6SF2*) gene was identified in the Dallas Heart Study cohort, in whom the *PNPLA3* I148M variant was first identified [38]. The 5′ end of *TM6SF2* is <5 kb downstream of the last exon in the *SUGP1* gene. Marker rs58542926 was not present on the genotyping chip we used for this study; however, it is in strong linkage disequilibrium (r^2=0.80) with rs10401969. The *TM6SF2* variant was also associated with significant reductions in plasma levels of alkaline phosphatase, triglycerides, and LDL-C, with heterozygous carriers manifesting intermediate levels; these results were replicated in two additional cohorts. In our analyses, we also found significant ($P < 0.0001$) differences in triglycerides and LDL-C for rs10401969 (Supplemental Table 2), although, in contrast with findings from the Dallas Heart Study cohort, minor allele homozygotes had higher mean levels than the heterozygous group. In addition, we found no statistically significant difference among genotype groups for alkaline phosphatase levels. TM6SF2 appears to play a role in lipoprotein synthesis and regulation of intestinal alkaline phosphatase activity, both of which may be impacted by extreme obesity and related co-morbidities that were not present in the Dallas Heart Study or the replication cohorts.

To evaluate other potential *NCAN* locus gene candidates, we systematically analyzed 20 genes located at this locus and found the only evidence for differential expression with *LPAR2*. This gene encodes the lysophosphatidic acid (LPA) receptor 2, which belongs to family I of the G-protein receptors and functions to mobilize calcium in response to LPA. Recently, LPA signaling was shown to impair glucose homeostasis and inhibit insulin secretion in obese mice fed a high-fat diet [39], and is involved with hepatic gluconeogenesis and insulin resistance [40]. Interestingly, the rs10401969 risk allele was associated with significantly higher glucose levels with a trend for higher insulin in the minor allele homozygotes (Supplemental Table 3). Notably, LPA acyltransferase activity was reduced 60%–70% with liver-specific *PNPLA3* knockdown [41], linking LPA, and potentially LPAR2, in a molecular pathway involving PNPLA3. These data support *LPAR2*

as a potential effector gene in the fatty liver *NCAN* locus. Thus, our data suggest that additional genes in the *SUGP1/TM6SF2/NCAN* locus may play a role in hepatic lipid accumulation in individuals with extreme obesity.

Consistent with other populations, we also found that grade of hepatic lipid content was related to several key metabolic parameters [42]. Mean levels of insulin, glucose, and HBA1c were normal in individuals without evidence of hepatic fat but increased even in grade 1, despite no changes in AST and ALT that would be suggestive of liver injury. Given that the mean BMI and waist circumference of individuals without hepatic fat was not different from those with high liver fat content and metabolic abnormalities, these levels of extreme obesity do not appear to play a role in the development of metabolic abnormalities. The increase in parameters of dysglycemia with hepatic fat content in the setting of extreme obesity provides circumstantial support for a causal role of hepatic lipid accumulation in the development of insulin resistance and subsequent type 2 diabetes [3].

Incipient fat accumulation, a phenotype seldom investigated, thus appears to be associated with metabolic abnormalities, although any genetic underpinnings have not yet been described. Our association analyses identified a SNP near the *SOCS2* gene that was protective against low grade hepatic fat accumulation. *SOCS2* was included in a core set of type 2 diabetes genes identified using a meta-analysis molecular network random sampling approach with data from multiple tissues in human and mouse [43]. *SOCS2* opposes the action of growth hormone, and deficiency of growth hormone results in hepatic lipid accumulation. Experimentally, *SOCS2* deficiency protects against hepatic lipid accumulation in a high fat diet-fed mouse model of fatty liver [44]. *SOCS2* was also upregulated in a chimeric mouse xenotransplant model containing human hepatocytes in which human growth hormone deficiency was found to cause liver lipid accumulation [45]. In these animals, treatment with human growth hormone improved hepatic lipid status, while significantly upregulating *SOCS2* expression. SOCS2 has also been shown to regulate proinsulin processing and insulin secretion in pancreatic beta cells, potentially linking fatty liver susceptibility and insulin resistance [46]. In a morbidly obese cohort of women, similar to the one studied here, *SOCS2* gene expression was decreased in liver RNA relative to patients who had lost a substantial amount of weight [47]. Our findings of reduced *SOCS2* expression in fatty compared to normal liver are concordant with these findings.

We also identified an association with low grade hepatic fat with a SNP nearby the *RAMP3* gene, which encodes a receptor activity modifying protein for the islet amyloid polypeptide (amylin) receptor. Amylin is a small peptide hormone expressed primarily by pancreatic islet beta cells [48]. Amylin improves the effect of leptin on insulin sensitivity in leptin-resistant diet-induced obese mice [49] and has been shown to regulate fatty acid sterification in cultured adipocytes [50]. The interaction of *SOCS2* and *RAMP3* genotypes, coupled with the available biological information, adds further evidence to support their role in mild hepatic fat accumulation.

Despite the significance of the findings reported here, we acknowledge several limitations. First, this study is comprised solely of individuals with extreme levels of obesity. The pathophysiology of hepatic lipid accumulation may be different in these extreme levels of

obesity. In regard to this potential limitation, however, two considerations are worth noting: 1) in our observations, the mean BMI at each successive grade of hepatic fat was similar, and 2) obtaining liver tissue from individuals without clinical indication, i.e., healthy subjects, presents a significant ethical challenge, given the potential morbidity and mortality associated with the procedure. Similarly, another potential limitation of this proposal is the relative uniformity of the study population, which is primarily of Caucasian ethnicity and largely represents a single institutional, retrospective design. The population is also over 80% female, which may also limit generalizability, although sex-specific analyses suggest that the *PNPLA3* locus may exhibit a higher degree of association in women than men. Our ability to detect significant associations was limited by the relatively small sample size for a GWAS. Unfortunately, large data sets of histologically based phenotypes for studies of human liver are difficult to accrue. Despite this limitation, we had sufficient power to replicate several associations at stringent statistical significance thresholds, largely due to the large effect sizes of the variants, e.g., those associated with total bilirubin and the *PNPLA3* SNPs. For subgroup analysis of incipient fat accumulation, which reduced the sample size, we estimated that we could detect a genotypic relative risk of about 1.3 with 80% power and a type I error rate of 0.05 given a risk allele frequency of 0.05 (e.g., the *RAMP3* SNP) and a prevalence of fatty liver of 0.62. This study represents the first large-scale genome-wide analysis of genetic variation and clinically relevant data on biopsy-proven fatty liver disease, and findings resulting from this work will be available for investigation in other ethnic groups and utilizing different study designs.

These findings replicate several loci for fatty liver-related phenotypes in extreme obesity and present evidence for new candidate genes that may play a role in the pathogenesis and metabolic consequences of hepatic fat accumulation. Further studies on the relationship between insulin resistance, incipient hepatic lipid loading, and the newly identified candidate genes may be warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the staff of the TGen Collaborative Sequencing Center for their assistance.

Funding

This work was supported by the National Institutes of Health DK091601 (JKD and GSG), P30 DK072488 (GSG and CDS), Translational Genomics Research Institute, and the Geisinger Obesity Institute.

Abbreviations

NAFLD	nonalcoholic fatty liver disease
PNPLA3	patatin-like phospholipase domain containing 3 gene
UGT1A	uridine 5′-diphospho-glucuronosyltransferase

SUGP1	SURP and G patch domain containing 1 gene
NCAN	neurocan
LPAR2	lysophosphatidic acid receptor 2
RAMP3	receptor (G protein-coupled) activity modifying protein 3
SOCS2	suppressor of cytokine signaling 2

References

1. Wang Y, Li YY, Nie YQ, Zhou YJ, Cao CY, Xu L. Association between metabolic syndrome and the development of non-alcoholic fatty liver disease. Exp Ther Med. 2013; 6:77–84. [PubMed: 23935723]
2. Fabbrini E, Sullivan S, Klein S. Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. Hepatology. 2010; 51:679–689. [PubMed: 20041406]
3. Birkenfeld AL, Shulman GI. Nonalcoholic fatty liver disease, hepatic insulin resistance, and type 2 Diabetes. Hepatology. 2014; 59:713–723. [PubMed: 23929732]
4. Day CP. Genes or environment to determine alcoholic liver disease and non-alcoholic fatty liver disease. Liver international: official journal of the International Association for the Study of the Liver. 2006; 26:1021–1028. [PubMed: 17032401]
5. Machado M, Marques-Vidal P, Cortez-Pinto H. Hepatic histology in obese patients undergoing bariatric surgery. J Hepatol. 2006; 45:600–606. [PubMed: 16899321]
6. Browning JD, Kumar KS, Saboorian MH, Thiele DL. Ethnic differences in the prevalence of cryptogenic cirrhosis. Am J Gastroenterol. 2004; 99:292–298. [PubMed: 15046220]
7. Caldwell SH, Harris DM, Patrie JT, Hespenheide EE. Is NASH underdiagnosed among African Americans? Am J Gastroenterol. 2002; 97:1496–1500. [PubMed: 12094872]
8. Struben VM, Hespenheide EE, Caldwell SH. Nonalcoholic steatohepatitis and cryptogenic cirrhosis within kindreds. The American journal of medicine. 2000; 108:9–13. [PubMed: 11059435]
9. Willner IR, Waters B, Patil SR, Reuben A, Morelli J, Riely CA. Ninety patients with nonalcoholic steatohepatitis: insulin resistance, familial tendency, and severity of disease. Am J Gastroenterol. 2001; 96:2957–2961. [PubMed: 11693332]
10. Schwimmer JB, Celedon MA, Lavine JE, et al. Heritability of nonalcoholic fatty liver disease. Gastroenterology. 2009; 136:1585–1592. [PubMed: 19208353]
11. Wilfred de Alwis NM, Day CP. Genes and nonalcoholic fatty liver disease. Current diabetes reports. 2008; 8:156–163. [PubMed: 18445359]
12. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nature genetics. 2008; 40:1461–1465. [PubMed: 18820647]
13. Hotta K, Yoneda M, Hyogo H, et al. Association of the rs738409 polymorphism in PNPLA3 with liver damage and the development of nonalcoholic fatty liver disease. BMC medical genetics. 2010; 11:172. [PubMed: 21176169]
14. Kantartzis K, Peter A, Machicao F, 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. [PubMed: 19651814]
15. Petit JM, Guiu B, Masson D, et al. PNPLA3 polymorphism influences liver fibrosis in unselected patients with type 2 diabetes. Liver international: official journal of the International Association for the Study of the Liver. 2011; 31:1332–1336. [PubMed: 21745307]
16. Romeo S, Sentinelli F, Dash S, 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. [PubMed: 19844213]

17. 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. [PubMed: 20684021]
18. Sookoian S, Castano GO, Burgueno AL, Gianotti TF, Rosselli MS, Pirola CJ. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. Journal of lipid research. 2009; 50:2111–2116. [PubMed: 19738004]
19. Wang CW, Lin HY, Shin SJ, et al. The PNPLA3 I148M polymorphism is associated with insulin resistance and nonalcoholic fatty liver disease in a normoglycaemic population. Liver international: official journal of the International Association for the Study of the Liver. 2011; 31:1326–1331. [PubMed: 21745282]
20. Liew PL, Lee WJ, Wang W, et al. Fatty liver disease: predictors of nonalcoholic steatohepatitis and gallbladder disease in morbid obesity. Obes Surg. 2008; 18:847–853. [PubMed: 18459024]
21. Chisholm J, Seki Y, Toouli J, Stahl J, Collins J, Kow L. Serologic predictors of nonalcoholic steatohepatitis in a population undergoing bariatric surgery. Surgery for obesity and related diseases: official journal of the American Society for Bariatric Surgery. 2012; 8:416–422. [PubMed: 21865094]
22. Dong H, Huang H, Yun X, et al. Bilirubin Increases Insulin Sensitivity in Leptin-Receptor Deficient and Diet-Induced Obese Mice Through Suppression of ER Stress and Chronic Inflammation. Endocrinology. 2014; 155:818–828. [PubMed: 24424052]
23. Cox AJ, Ng MC, Xu J, et al. Association of SNPs in the UGT1A gene cluster with total bilirubin and mortality in the Diabetes Heart Study. Atherosclerosis. 2013; 229:155–160. [PubMed: 23642732]
24. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005; 41:1313–1321. [PubMed: 15915461]
25. Wood GC, Chu X, Manney C, et al. An electronic health record-enabled obesity database. BMC medical informatics and decision making. 2012; 12:45. [PubMed: 22640398]
26. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008; 24:2938–2939. [PubMed: 18974171]
27. Still CD, Craig Wood G, Chu X, et al. Clinical factors associated with weight loss outcomes after Roux-en-Y gastric bypass surgery. Obesity (Silver Spring). 2013
28. Speliotes EK, Yerges-Armstrong LM, Wu J, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS genetics. 2011; 7:e1001324. [PubMed: 21423719]
29. Dai X, Wu C, He Y, et al. A genome-wide association study for serum bilirubin levels and gene-environment interaction in a Chinese population. Genetic epidemiology. 2013; 37:293–300. [PubMed: 23371916]
30. Chen G, Ramos E, Adeyemo A, et al. UGT1A1 is a major locus influencing bilirubin levels in African Americans. European journal of human genetics: EJHG. 2012; 20:463–468. [PubMed: 22085899]
31. Johnson AD, Kavousi M, Smith AV, et al. Genome-wide association meta-analysis for total serum bilirubin levels. Human molecular genetics. 2009; 18:2700–2710. [PubMed: 19414484]
32. Kitamoto T, Kitamoto A, Yoneda M, et al. Genome-wide scan revealed that polymorphisms in the PNPLA3, SAMM50, and PARVB genes are associated with development and progression of nonalcoholic fatty liver disease in Japan. Human genetics. 2013; 132:783–792. [PubMed: 23535911]
33. Chalasani N, Guo X, Loomba R, et al. Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. Gastroenterology. 2010; 139:1567–1576. 1576 e1561–1566. [PubMed: 20708005]
34. Speliotes EK, Butler JL, Palmer CD, et al. PNPLA3 variants specifically confer increased risk for histologic nonalcoholic fatty liver disease but not metabolic disease. Hepatology. 2010; 52:904–912. [PubMed: 20648472]

35. Gorden A, Yang R, Yerges-Armstrong LM, et al. Genetic variation at NCAN locus is associated with inflammation and fibrosis in non-alcoholic fatty liver disease in morbid obesity. Human heredity. 2013; 75:34–43. [PubMed: 23594525]
36. Verrijken A, Beckers S, Francque S, et al. A gene variant of PNPLA3, but not of APOC3, is associated with histological parameters of NAFLD in an obese population. Obesity. 2013; 21:2138–2145. [PubMed: 23512881]
37. Gordon L, Joo JE, Powell JE, et al. Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. Genome Res. 2012; 22:1395–1406. [PubMed: 22800725]
38. Kozlitina J, Smagris E, Stender S, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nature genetics. 2014; 46:352–356. [PubMed: 24531328]
39. Rancoule C, Attane C, Gres S, et al. Lysophosphatidic acid impairs glucose homeostasis and inhibits insulin secretion in high-fat diet obese mice. Diabetologia. 2013; 56:1394–1402. [PubMed: 23508306]
40. Sankella S, Garg A, Horton JD, Agarwal AK. Hepatic gluconeogenesis is enhanced by phosphatidic acid which remains uninhibited by insulin in lipodystrophic Agpat2−/− mice. The Journal of biological chemistry. 2014; 289:4762–4777. [PubMed: 24425876]
41. Kumashiro N, Yoshimura T, Cantley JL, et al. Role of patatin-like phospholipase domain-containing 3 on lipid-induced hepatic steatosis and insulin resistance in rats. Hepatology. 2013; 57:1763–1772. [PubMed: 23175050]
42. Kashyap SR, Diab DL, Baker AR, et al. Triglyceride levels and not adipokine concentrations are closely related to severity of nonalcoholic fatty liver disease in an obesity surgery cohort. Obesity (Silver Spring). 2009; 17:1696–1701. [PubMed: 19360015]
43. Rasche A, Al-Hasani H, Herwig R. Meta-analysis approach identifies candidate genes and associated molecular networks for type-2 diabetes mellitus. BMC genomics. 2008; 9:310. [PubMed: 18590522]
44. Zadjali F, Santana-Farre R, Vesterlund M, et al. SOCS2 deletion protects against hepatic steatosis but worsens insulin resistance in high-fat-diet-fed mice. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2012; 26:3282–3291. [PubMed: 22562833]
45. Tateno C, Kataoka M, Utoh R, et al. Growth hormone-dependent pathogenesis of human hepatic steatosis in a novel mouse model bearing a human hepatocyte-repopulated liver. Endocrinology. 2011; 152:1479–1491. [PubMed: 21303949]
46. Lebrun P, Cognard E, Gontard P, et al. The suppressor of cytokine signalling 2 (SOCS2) is a key repressor of insulin secretion. Diabetologia. 2010; 53:1935–1946. [PubMed: 20499047]
47. Elam MB, Cowan GS Jr, Rooney RJ, et al. Hepatic gene expression in morbidly obese women: implications for disease susceptibility. Obesity. 2009; 17:1563–1573. [PubMed: 19265796]
48. Westermark GT, Westermark P. Islet amyloid polypeptide and diabetes. Current protein & peptide science. 2013; 14:330–337. [PubMed: 23745697]
49. Lutz TA. The role of amylin in the control of energy homeostasis. American journal of physiology Regulatory, integrative and comparative physiology. 2010; 298:R1475–1484.
50. Ye JM, Lim-Fraser M, Cooney GJ, et al. Evidence that amylin stimulates lipolysis in vivo: a possible mediator of induced insulin resistance. American journal of physiology Endocrinology and metabolism. 2001; 280:E562–569. [PubMed: 11254462]

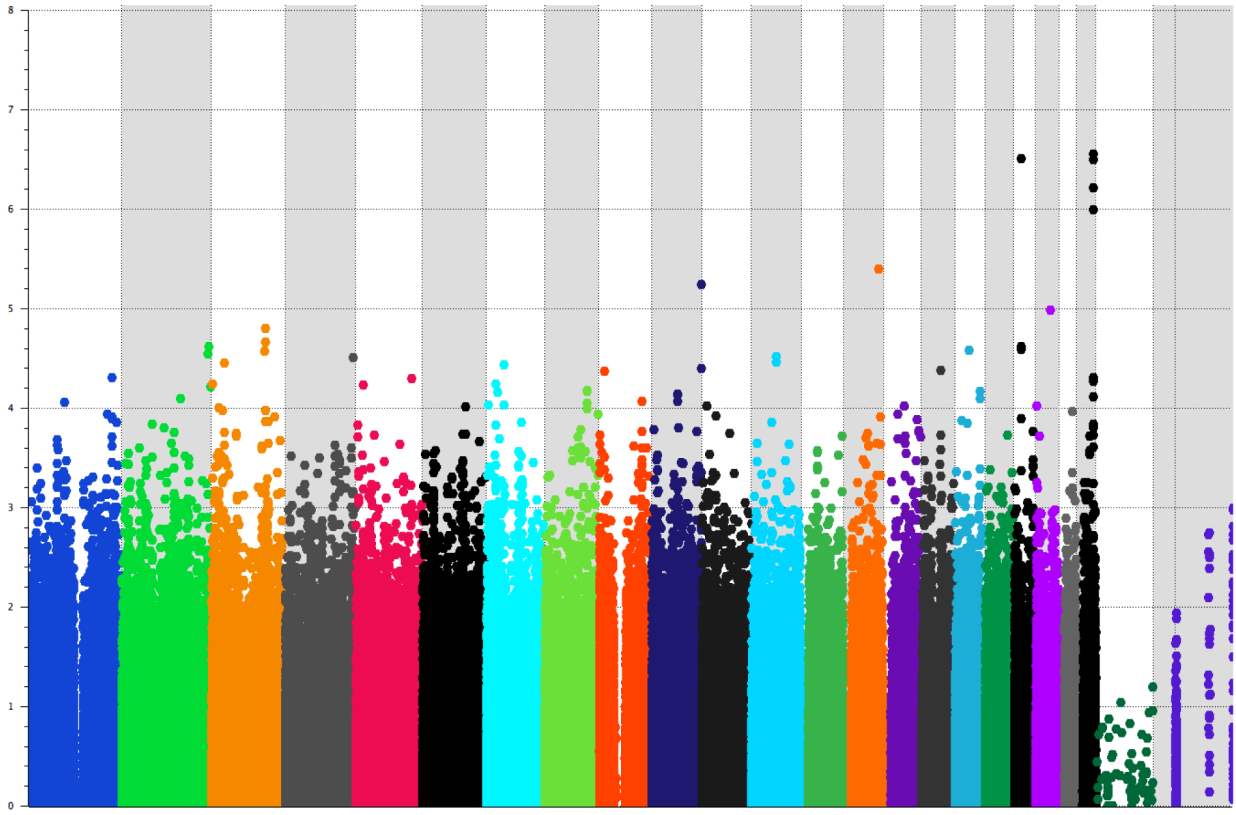

Figure 1.
Linear regression association analysis of hepatic lipid grade using an additive genetic model.

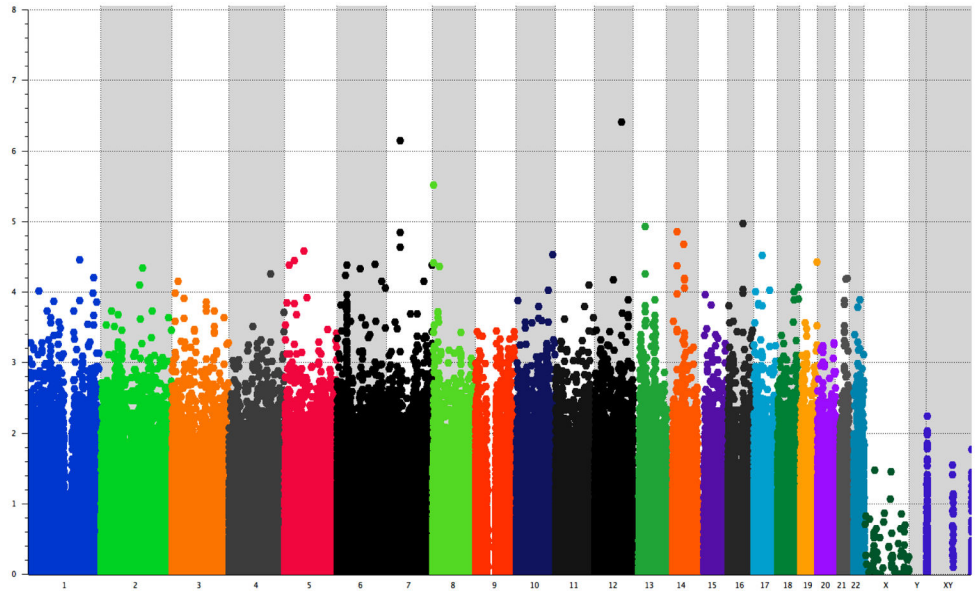

Figure 2.

Logistic regression association analysis of hepatic lipid grade 1 versus grade 0 using an additive genetic model.

Table 1

Cohort characteristics by grade of hepatic fat content

Variable (n)	ALL* (1868)	S.D.	GRADE 0* (482)	S.D.	GRADE 1* (798)	S.D.	GRADE 2* (408)	S.D.	GRADE 3* (180)	S.D.	P VALUE (ANOVA)
Female %(n)	81 (1512)		86 (416)		81 (645)		77 (312)		78 (139)		0.0012**
Male %(n)	19 (356)		14 (66)		19 (153)		23 (96)		22 (41)		
BMI (kg/m2)	47.0	8.2	46.2	8.0	47.0	8.0	47.6	8.5	48.1	9.2	0.019
Waist Circumference (in)	53.6	6.7	52.2	6.4	53.8	6.8	54.4	6.8	54.3	6.3	<0.0001
Age (Years)	45	11	45	12	45	11	45	11	46	9	0.477
Glucose (mg/dL)	108	46	96	27	105	40	125	65	117	46	<0.0001
Insulin (IU)	25	28	18	20	24	27	32	35	33	25	<0.0001
HBa1c (%)	6.3	1.3	6.0	1.1	6.2	1.3	6.7	1.5	6.8	1.5	<0.0001
Triglycerides (mg/dL)	170	111	139	66	166	92	202	159	204	126	<0.0001
Total Cholesterol (mg/dL)	185	40	184	38	186	39	187	42	186	40	0.652
HDL (mg/dL)	46	11	50	12	46	11	44	11	45	11	<0.0001
LDL (mg/dL)	106	34	107	34	107	33	104	34	102	34	0.212
Chol/HDL Ratio	4.2	1.2	3.9	1.1	4.2	1.1	4.4	1.3	4.4	1.3	<0.0001
ALT (U)	30	18	24	12	28	16	36	19	46	27	<0.0001
AST (U)	26	12	23	8	25	10	29	12	36	19	<0.0001
Alkaline Phosphatase (U)	80	24	80	22	81	26	80	23	77	20	0.232
Total Bilirubin (mg/dL)	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.699

* Mean values

** Chi-square

Table 2

Association of variants with total bilirubin, hepatic fat grade, and incipient fat accumulation

Phenotype	SNP	Locus	P-value		FDR P-value		Bonf. P-value	
			Add	Dom	Add	Dom	Add	Dom
Total bilirubin	rs4148325	*UGT1A1*	5×10^{-93}		3×10^{-87}		3×10^{-87}	
	rs6742078	*UGT1A1*	1×10^{-92}		4×10^{-87}		8×10^{-87}	
	rs887829	*UGT1A1*	2×10^{-92}		3×10^{-87}		1×10^{-86}	
	rs4148324	*UGT1A1*	2×10^{-92}		4×10^{-87}		$1X\ 10^{-86}$	
Hepatic Fat Grades 0–3	rs10401969	*SUGP1*	3.1×10^{-7}	5.6×10^{-7}	0.09	0.085	0.19	0.34
	rs4823173	*PNPLA3*	2.8×10^{-7}	7.9×10^{-8}	0.17	0.048	0.17	0.048
	rs2896019	*PNPLA3*	3.1×10^{-7}	1.1×10^{-7}	0.06	0.034	0.19	0.068
	rs2281135	*PNPLA3*	6.0×10^{-7}	2.4×10^{-7}	0.09	0.049	0.36	0.15
Hepatic Fat Grade 1 vs. 0	rs10859525	*SOCS2*	3.8×10^{-7}		0.23		0.23	
	rs1294908	*RAMP3*	7.1×10^{-7}		0.21		0.43	