The Genetic Epidemiology of Nonalcoholic Fatty Liver Disease

Toward a Personalized Medicine

Silvia Sookoian, MD, PhD[a,*], Carlos J. Pirola, PhD[b,*]

KEYWORDS

• Nonalcoholic fatty liver disease • Nonalcoholic steatohepatitis • Adiponutrin
• PNPLA3 • Epigenetics • Systems biology • Genetics • Drug targets

KEY POINTS

- Nonalcoholic fatty liver disease (NAFLD) is a complex disorder that develops from the interplay of a myriad of genetic and environmental factors.
- Until some years ago, the most helpful strategy in the search for genes underlying complex diseases such as NAFLD was to look at candidate genes.
- The knowledge of the genetic bases of NAFLD has tremendously benefited from recent advances in genotyping technology and information generated by genome-wide association studies.
- Now, systems biology offers the opportunity to get more insight in its pathophysiology and opens new avenues for novel treatments.

INTRODUCTION

This review, very far from the collected and summarized information about the current knowledge of the genetic epidemiology of nonalcoholic fatty liver disease (NAFLD), attempts to build new integrated ways to understand the genetic architecture of the

Funding: This study was partially supported by grants PICT 2008-1521 and PICT 2010-0441 (Agencia Nacional de Promoción Científica y Tecnológica), and UBACYT CM04 (Universidad de Buenos Aires). S.S. and C.J.P. belong to Consejo Nacional de Investigaciones Científicas (CONICET).
Disclosures: None.
[a] Department of Clinical and Molecular Hepatology, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires–National Council of Scientific and Technological Research (CONICET), Combatientes de Malvinas 3150 Ciudad Autónoma de Buenos Aires-1427, Argentina; [b] Department of Molecular Genetics and Biology of Complex Diseases, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires–National-Council of Scientific and Technological Research (CONICET), Combatientes de Malvinas 3150 Ciudad Autónoma de Buenos Aires-1427, Argentina
* Corresponding authors. Instituto de Investigaciones Médicas A. Lanari-IDM CONICET, Combatiente de Malvinas 3150, Buenos Aires (1427), Argentina.
E-mail addresses: sookoian.silvia@lanari.fmed.uba.ar; pirola.carlos@lanari.fmed.uba.ar

Clin Liver Dis 16 (2012) 467–485
doi:10.1016/j.cld.2012.05.011 **liver.theclinics.com**
1089-3261/12/$ – see front matter

disease. Hence, based on current knowledge about a disease-associated *PNPLA3* gene variant, the authors postulate new paradigms of NAFLD pathogenesis. Also proposed is a systems biology approach in the search of joining effects of genetic variants from multiple pathways. Based on the information of previously reported loci associated with the genetic risk of NAFLD, candidate genes from the entire genome are prioritized to show putative unexplored genes as potential modifiers of the biology of the disease. An important question is whether it is possible to rapidly translate the knowledge about genetic susceptibility of NAFLD into more individualized decision making and personalized medicine. If so, can the diagnostic and therapeutic strategies be improved? This article attempts to answer whether disease-associated gene variants can help select therapeutic targets. Therefore, an effort is made to translate the current findings of NAFLD-associated gene-risk variants into therapeutic target identification. Finally, the article further explores DNA sequence variation to introduce relevant knowledge about the role of epigenetics in NAFLD.

Nonalcoholic fatty liver disease (NAFLD) is a complex disorder that develops from the interplay of a myriad of genetic and environmental factors. This concept clearly characterizes the nature of the disease as highly heterogeneous and influenced by multiple critical pathways. The challenge, as well the prospect of the genomic era, lies in the ability to integrate all the information about the genetic components of a disease to better understand its pathogenesis as well as its treatment strategies.

Until some years ago, the most helpful strategy in the search for genes underlying complex diseases such as NAFLD was to look at candidate genes. The candidate genes for NAFLD were then selected based either on their known or presumed function or on their biological plausibility in the disease pathophysiology; a recent comprehensive review summarizes the major findings on candidate-gene association studies in NAFLD.[1] Most candidate-gene association studies showed a small effect on the susceptibility of developing fatty liver, and a few of them demonstrated an effect on disease progression.[2–5]

The knowledge of the genetic bases of NAFLD has tremendously benefited from recent advances in genotyping technology and information generated by genome-wide association studies (GWAS). In fact, the first GWAS on NAFLD[6] has not only expanded our knowledge about the genetic component of the disease but has also contributed to the understanding about the role of genes on the histologic severity.[7] In this particular GWAS, differing from the most common genome scans, Romeo and colleagues[6] specifically surveyed nonsynonymous sequence variations along the entire genome. As a result, the missense rs738409 C/G single-nucleotide polymorphism (SNP) implying an amino acid change from isoleucine (I) to methionine (M) at the position 148 (I148M) of the protein encoding by the patatin-like phospholipase domain–containing 3 gene (*PNPLA3*), also known as adiponutrin, was strongly associated with increased hepatic fat content measured by ^{1}H magnetic resonance spectroscopy. After further replication in several populations around the world, including adults and children, the G allele in the forward strand was significantly and unequivocally associated with increased risk of hepatic triglyceride accumulation and fatty liver disease, as shown in a recent meta-analysis.[7] In fact, carriers of the GG homozygous genotype show 73% higher lipid fat content than do carriers of the CC genotype. Surprisingly, even though ethnic differences in the susceptibility to NAFLD are evident,[8] the effect of the *PNPLA3* variant transcends racial or ethnic boundaries, as Asian and Caucasians have an almost identical effect (odds ratio [OR] 3.26 vs 3.11). A significantly higher effect was demonstrated in females, however, than in males, a sexual dimorphism discussed later.[7] In conclusion, the risk effect of the rs738409 on developing fatty liver is perhaps one of the strongest ever reported for

a common variant modifying the genetic susceptibility for complex diseases (explaining 5.3% of the total variance), and seems to follow an additive model, at least for liver fatty accumulation.[7]

Moreover, the rs738409 is nowadays the most replicated gene variant in the study of the genetic component of disease severity as, after pooling liver biopsies of 2124 NAFLD patients, nonalcoholic steatohepatitis (NASH) was more frequently observed in GG than in CC homozygous carriers (OR 3.488, 95% confidence interval [CI] 1.859–6.545; data from 2124 patients).[7] Even more surprising is that a large body of literature about replication studies of *PNPLA3* and NAFLD is constantly generated, with more articles published annually than the entire evidence accumulated until now about candidate-gene association studies for this disease. Thus, the enthusiasm about newly discovered loci has vanished and the newly discovered rs738409 SNP far more widely covered than the previously proposed candidate genes.

Although no doubts exist about the impact of the rs738409 on the natural history of NAFLD, several important questions still remain unanswered and some other questions could be resolved, based on the current knowledge about the impact of the *PNPLA3* variant on NAFLD; this review introduces some ideas on this matter.

PNPLA3 AND NAFLD: FROM GENES TO DISEASE PATHOGENESIS. *PNPLA3* AND THE LIVER "FAT REMODELING" HYPOTHESIS

Considering that NAFLD is a polygenic and multifactorial disease that results from complex interactions among multiple genetic factors in addition to a collection of environmental exposures, it makes probably no sense to attribute to *PNPLA3* the whole burden of the genetic risk of the disease. Thus, the first question to arise after the findings of the NAFLD GWAS is how a single variant in a single gene is able to cause, by itself, such a strong impact on the biology of the disease. Certainly SNPs that alter the coding sequence and result in a nonsynonymous change have more chance of leading to a pronounced effect in the protein function that markedly affects the disease or trait of interest. That would be case for the rs738409, which involves a coding variant that encodes the amino acid substitution I148M; this change in the amino acid sequence seems to cause a loss of protein function, altering the hydrolysis of glycerolipids, more popularly known as fatty acid triesters of glycerol or triacylglycerols (TAG).[9]

One may wonder whether the disruption of the hydrolysis of glycerolipids is so important in the context of NAFLD. Previous evidence showed that defects or perturbations in the hydrolysis of glycerolipids (both neutral and phospholipids) are associated with the cluster of metabolic disorders, characteristics of the metabolic syndrome, including type 2 diabetes, obesity, and NAFLD.[10]

The hydrolysis of TAG to diacylglycerol (DAG) is primarily achieved by the enzyme PNPLA2 and putatively by PNPLA3.[11] In fact, there is a group of genes (PNPLA1 to PNPLA9) that encode for proteins (enzymes) containing a patatin-like domain with broad lipid acyl-hydrolase activity and with specificities for diverse substrates such as triacylglycerols, phospholipids, and retinol esters.[12] In this regard the strong effect of the rs738409 variant is surprising, considering that the PNPLA3 protein shares domains with several family members and is coexpressed with PTGES (prostaglandin E synthase). In silico function prediction of the *PNPLA3* gene[13] shows that only PNPLA2, PNPLA3, PNPLA8, and PLA2G6 have carboxylesterase and lipase activities ($P<4.8$ and 9.4×10^{-5}) and PNPLA3, PNPLA8, and PLA2G6 have phospholipase A2 activity ($P<1.6 \times 10^{-4}$), indicating that the release of arachidonic acid and its product, prostaglandin E_2, may have a role in the pathogenesis of NAFLD. Supporting this observation, Puri and colleagues[14] recently showed that NASH is associated with

increased levels of the proinflammatory product 15-hydroxyeicosatetraenoic acid (HETE) and 11-HETE, a nonenzymatic oxidation product of arachidonic acid.

In addition, neutral lipid, triglyceride, or glycerolipid catabolic processes are shared by only PNPLA2 and PNPLA3 ($P<.05$). Although these aspects deserve further investigation, the functional redundancy explained previously can also explain the paradoxic results observed in mice with global targeted deletion of *Pnpla3*, which do not reproduce the biological effect reported in humans.[15,16]

The hallmark feature of NAFLD is the abnormal accumulation of TAG and DAG in the hepatocytes. The evidence showed that stored hepatic TAG are largely hydrolyzed to DAG and then reesterified before being secreted as very-low-density lipoprotein TAG.[17] In this process, there is a critical step of remodeling of DAGs, which are reesterified to TAG before being secreted. Overall, these data may support the hypothesis proposed by Jenkins and colleagues[18] suggesting that PNPLA3 might stimulate triacylglycerol/fatty acid cycling (fatty acid reesterification). Taking into account this evidence, it is reasonable to speculate that PNPLA3 plays a role in hepatic lipid partitioning while also bearing in mind that DAG is a potent activator of protein kinases C (PKCs) with potential metabolic effects per se, as activation of some conventional and novel PKCs in response to increased levels of DAG have been shown to counteract insulin signaling.[19] Hence, in this scenario and considering the strong impact of rs738409 on the natural history of NAFLD, the contribution of the knowledge of the disease-associated genes, at least but not last, leads one to advice about changing the paradigm of the NAFLD pathogenesis and starting to revise the "2-hit hypothesis" to explain the molecular events triggering the disease. Supporting this notion, recent evidence clearly showed that changes in lipid partitioning in the liver are associated with an increase in liver cell apoptosis,[20] a common finding in patients and rodent models of NAFLD.[21,22]

PNPLA3 AND GENE BY SEX INTERACTION: DOES THE RS738409 VARIANT EXPLAIN SEXUAL DIMORPHISM OF NAFLD?

As a final comment, the findings about the rs738409 brought new answers to poorly explored questions, for instance, the role of sexual dimorphism in the pathogenesis of NAFLD. As already mentioned, meta-regression analysis of pooled data from all the published evidence showed a negative association between the effect of rs738409 on liver fat content and the proportion of male individuals in the studied populations.[7] Despite the explanation of this gene by sex interaction still being unknown, one may speculate that the effect of *PNPLA3* and, putatively, the rs738409 risk variant may be modulated by sex hormones. This explanation is in agreement with previous evidence that PNPLA3 levels are strongly influenced by metabolic status and hormones, such as insulin.[23] Reinforcing the previous concept about the role of PNPLA3 on lipid remodeling, sexual hormones, such as estrogen, modulate lipogenic genes, including *SREBP-1c*, and participate in adiposity and fuel partitioning.[24]

NAFLD AND GENETIC SUSCEPTIBILITY: SYSTEMS BIOLOGY APPROACHES AND THE SEARCH OF JOINING EFFECTS OF GENETIC VARIANTS FROM MULTIPLE PATHWAYS

As summarized earlier, several SNPs in different genes or loci have been proposed as potential modifiers (albeit modest) of the genetic risk of NAFLD, even though the mechanism behind the association between the associated gene variant and the disease may be unknown. Instead of understanding the effect of each reported variant independently, one should consider integration of the current knowledge about the genetic influence of NAFLD into common pathways of the disease, and ask whether

common regulatory pathways or common physiologic processes link the pathophysiology of NAFLD with the metabolic syndrome.

To answer these questions, the authors used systems biology approaches to integrate genomic, molecular, and physiologic data to decipher putative pathways that connect all the available information about genes suspected to play a role in the susceptibility of NAFLD. This approach is designed to analyze and integrate genomic, transcriptomic, and/or proteomic data to inferred from genetic signals–related pathways of disease. Furthermore, systems biology introduces a new concept for revealing the pathogenesis of human disorders, and suggests that the presence of common physiologic processes and molecular networks influences the risk of a disease.

Based on this hypothesis, different systems biology approaches were proposed, such as gene-enrichment analysis and the use of a protein-protein interaction network.

For this purpose, the authors built a candidate gene list using as template the published evidence about the genetic component of NAFLD from all the candidate-gene association studies, including *PNPLA3* (the input gene list is given in **Table 1**). The list includes 58 reported loci with variants associated with either fatty liver or NASH, as mentioned previously.[1]

Functional enrichment analysis of the gene list was performed by the bioinformatic resource ToppGene Suite (http://toppgene.cchmc.org) and ToppCluster (http://toppcluster.cchmc.org/), based on Transcriptome, Proteome, Regulome (TFBS and miRNA), Ontologies (GO, Pathway), Phenotype (human disease and mouse phenotype), Pharmacome (Drug-Gene associations), literature cocitation, and other features. The analysis showed that the 58 reported loci could be integrated into several common functional pathways, of which the highly ranked are shown in **Fig. 1**. The predicted pathways are mainly enriched with mechanisms of cellular control of lipid and lipoprotein metabolism. Thus, the associated genes strongly suggest the role of lipotoxicity in the pathogenesis of NAFLD. Moreover, as shown in **Table 2**, highly predicted biological processes were observed, most of them related to the regulation of lipid homeostasis and the cellular lipid metabolic process, and whereby *PNPLA3* was jointly included with other related genes.

Gene-enrichment analysis also detected common physiologic processes that potentially link the pathophysiology of NAFLD with the metabolic syndrome; for example, 3 pathways were predicted that are involved in cardiovascular system regulation (see **Table 2**).

In addition, a functional association analysis was performed that included protein and genetic interactions, pathways, coexpression, colocalization, and protein domain similarity, using the bioinformatic resource GenMANIA.[13] Several genes are regarded as direct "neighbors" of the *PNPLA3* considering coexpression or colocalization, physical interactions, belonging to the same pathways, or having shared protein domains (**Fig. 2**). Most of them are not obvious *PNPLA3*-related genes, but then again, some interesting associations with prostaglandin-endoperoxide synthases (ie, *PTGS2*) emerged that may explain the beneficial effect of indomethacin on the liver fat accumulation observed in an experimental rodent model of NAFLD.[25]

GENE PRIORITIZATION BASED ON PREVIOUS REPORTED LOCI SHOWS NUCLEAR RECEPTORS AND HYPOXIA AS MOLECULAR MEDIATORS OF NAFLD PATHOGENESIS

As already mentioned, the enthusiasm about the discovery of novel loci associated with NAFLD is nowadays vanishing, and this could have a negative impact on our ability

Table 1
Candidate gene list based on the published evidence about the genetic component of NAFLD

Gene Symbol	Ensembl ID	Gene Name Description
PEMT	ENSG00000133027	Phosphatidylethanolamine *N*-methyltransferase
MTTP	ENSG00000138823	Microsomal triglyceride transfer protein large subunit precursor
APOC3	ENSG00000110245	Apolipoprotein C-III precursor (Apo-CIII)
NR1I2	ENSG00000144852	Orphan nuclear receptor PXR (pregnane X receptor)
FABP2	ENSG00000145384	Fatty acid-binding protein, intestinal
DGAT1	ENSG00000185000	Diacylglycerol *O*-acyltransferase 1
DGAT2	ENSG00000062282	Diacylglycerol *O*-acyltransferase 2
ACSL4	ENSG00000068366	Long-chain-fatty-acid–CoA ligase 4
ADRB3	ENSG00000188778	β3-Adrenergic receptor (β3 adrenoceptor)
ADRB2	ENSG00000169252	β2-Adrenergic receptor (β2 adrenoceptor)
LIPC	ENSG00000166035	Hepatic triacylglycerol lipase precursor
APOE	ENSG00000130203	Apolipoprotein E precursor (Apo-E)
CLOCK	ENSG00000134852	Circadian locomotor output cycles protein kaput
ENPP1	ENSG00000197594	Ectonucleotide pyrophosphatase/phosphodiesterase 1
IRS1	ENSG00000169047	Insulin receptor substrate 1
ADIPOQ	ENSG00000181092	Adiponectin precursor
ADIPOR1	ENSG00000159346	Adiponectin receptor protein 1
ADIPOR2	ENSG00000006831	Adiponectin receptor protein 2
PPARA	ENSG00000186951	Peroxisome proliferator-activated receptor α
PPARG	ENSG00000132170	Peroxisome proliferator-activated receptor γ
PPARGC1A	ENSG00000109819	Peroxisome proliferator-activated receptor γ coactivator 1α
TCF7L2	ENSG00000148737	Transcription factor 7-like 2
GCKR	ENSG00000084734	Glucokinase regulatory protein
MC4R	ENSG00000166603	Melanocortin receptor 4
*SPINK*1	ENSG00000164266	Pancreatic secretory trypsin inhibitor precursor
LEPR	ENSG00000116678	Leptin receptor precursor
TNF	ENSG00000204490	Tumor necrosis factor precursor (TNF-α)
TNFSF10	ENSG00000121858	Tumor necrosis factor ligand superfamily member 10
IL6	ENSG00000136244	Interleukin-6 precursor
CD14	ENSG00000170458	Monocyte differentiation antigen CD14 precursor
GCLC	ENSG00000001084	Glutamate–cysteine ligase catalytic subunit
SOD2	ENSG00000112096	Superoxide dismutase [Mn], mitochondrial precursor
HFE	ENSG00000010704	Hereditary hemochromatosis protein precursor (HLA-H)
UGT1A1	ENSG00000167165	UDP-glucuronosyltransferase 1–8 precursor
UCP1	ENSG00000109424	Mitochondrial brown fat uncoupling protein 1
PTGS2	ENSG00000073756	Prostaglandin G/H synthase 2 precursor
ABCB11	ENSG00000073734	Bile salt export pump (ATP-binding cassette subfamily B member 11)
MIF	ENSG00000099964	Macrophage migration inhibitory factor
CYP2E1	ENSG00000130649	Cytochrome P450 2E1
SERPINA1	ENSG00000197249	α1-Antitrypsin precursor

(*continued on next page*)

Table 1
(*continued*)

Gene Symbol	Ensembl ID	Gene Name Description
MTHFR	ENSG00000177000	Methylene tetrahydrofolate reductase
IL1B	ENSG00000125538	Interleukin-1β precursor
TLR4	ENSG00000136869	Toll-like receptor 4 precursor
CFTR	ENSG00000001626	Cystic fibrosis transmembrane conductance regulator
STAT3	ENSG00000168610	Signal transducer and activator of transcription 3
AGTR1	ENSG00000144891	Type-1 angiotensin II receptor
KLF6	ENSG00000067082	Krueppel-like factor 6
PNPLA3	ENSG00000100344	Adiponutrin (iPLA2-ε) (calcium-independent phospholipase A2ε) (Patatin-like phospholipase domain-containing protein 3)
FDFT1	ENSG00000079459	Squalene synthetase
COL13A1	ENSG00000197467	α1 type XIII collagen isoform 1
PDGFA	ENSG00000197461	Platelet-derived growth factor A chain precursor
LTBP3	ENSG00000168056	Latent-transforming growth factor β–binding protein 3 precursor
EFCAB4B	ENSG00000130038	EF-hand calcium binding domain 4B
NCAN	ENSG00000130287	Neurocan core protein precursor
LYPLAL1	ENSG00000143353	Lysophospholipase-like protein 1
GCKR	ENSG00000084734	Glucokinase (hexokinase 4) regulator 1
PPP1R3B	ENSG00000173281	Protein phosphatase 1, regulatory (inhibitor) subunit 3B

to better understand the disease pathogenesis. Hence, the already known associated disease loci might serve as a template to look further for gene variants involved in the disease susceptibility. This approach benefits from examining previously proposed candidate gene loci in a more systematic manner. For this reason, the authors performed a comprehensive analysis of candidate regions generated by the freely accessible ENDEAVOUR software available at http://homes.esat.kuleuven.be/~bioiuser/endeavor/endeavor.php, which specifically focused on the previous loci reported to be associated with NAFLD susceptibility (see **Table 1**).

ENDEAVOUR is a software application for the computational prioritization of candidate genes underlying biological processes or diseases, based on their similarity to known genes involved in a disease as previously described.[26] The hypothesis of prioritization by ENDEAVOUR is that candidate test genes are ranked based on their similarity with a set of known training genes. This attractive methodology allows performance of gene prioritization based on the biological plausibility of a gene-disease association and on the knowledge of the protein function. In addition, this strategy allows expansion of the selection of putative candidate genes and prediction of new targets, as the authors have reported previously for other components of the metabolic syndrome.[27] Surprisingly, a prioritized list of candidate genes, along with the rankings per data source, showed interesting loci never previously explored that could be tested for gene-associated risk variants (**Fig. 3**). For instance, the results of gene prioritization ranked very highly 2 nuclear receptors, the *NR1H4* or farnesoid X nuclear receptor, a ligand-activated transcription factor that functions as a receptor for bile acids such as chenodeoxycholic acid, lithocholic acid, and deoxycholic acid, and the *RXRA* or retinoid X receptor, a nuclear receptor that mediates the biological

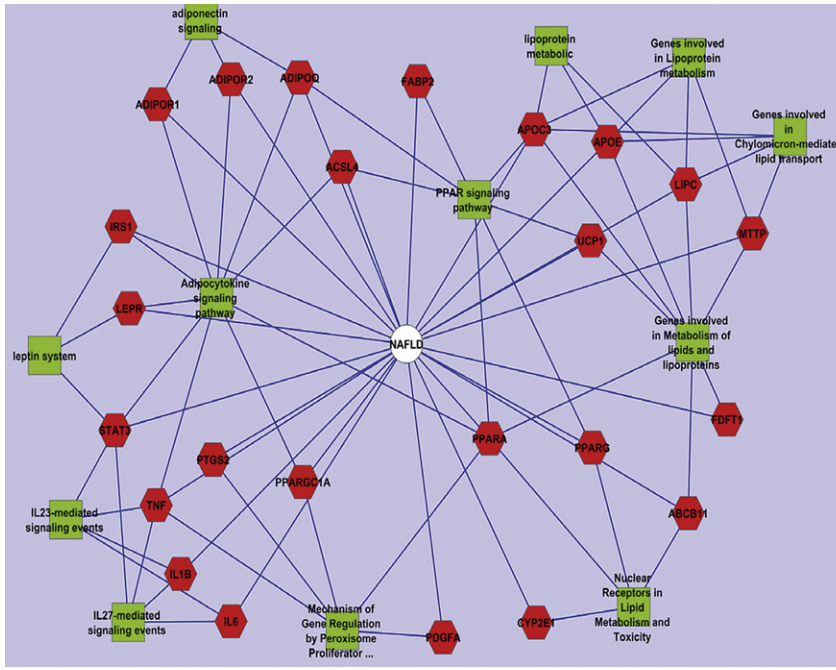

Fig. 1. Functional enrichment analysis of loci previously associated with NAFLD: predicted disease pathways. Disease pathways were obtained by the Toppcluster tool, and the analysis was based on combined data from the following resources: BioCyc, GenMAPP, KEGG pathway, MSigDB: C2.cp—BioCarta, MSigDB: C2.cp—Reactome, 430MSigDB: C2.cp—Sigma-Aldrich, MSigDB: C2.cp—Signaling Gateway, MSigDB: C2.cp—Signaling Transduction KE, MSigDB: C2.cp—SuperArray, NCI-Nature Curated, PantherDB, Pathway Ontology, SMPDB. List of pathway IDs in the figure: Adipocytokine signaling pathway, hsa04920; Adiponectin signaling, PW:0000563; PPAR signaling pathway, hsa03320; Nuclear receptors in lipid metabolism and toxicity, BIOCARTA NUCLEAR RS PATHWAY; Genes involved in chylomicron-mediated lipid transport, REACTOME CHYLOMICRON-MEDIATED LIPID TRANSPORT; Genes involved in the metabolism of lipids and lipoproteins, REACTOME METABOLISM OF LIPIDS AND LIPOPROTEINS; IL27-mediated signaling events, il27pathway; Genes involved in lipoprotein metabolism, REACTOME LIPOPROTEIN METABOLISM; Mechanism of gene regulation by peroxisome proliferators via PPARα, BIOCARTA PPARA PATHWAY; Leptin system, PW:0000363; Lipoprotein metabolic, PW:0000482; IL23-mediated signaling events, il23pathway.

effects of retinoids by their involvement in retinoic acid-mediated gene activation. This information is a reminder of the important role the bile acids might have in the regulation of lipid homeostasis,[28] and suggests that these candidate genes may be attractive pharmacologic targets. It should be emphasized that bile acids may act through the elevation of incretins,[29] and this family of substances has been proved to be effective in improving NAFLD.[30]

Moreover, the data mining also identified *HIF1A* (hypoxia-inducible factor 1, α subunit), a transcription factor involved in the regulation of cellular response to hypoxia. Hypoxia is a pathogenic disease pathway largely ignored in NAFLD, but is now recognized as a putative mediator of initial metabolic changes in the liver[21,22] and in the progression of inflammation and liver damage.[31,32] Other examples that are worth noting are *HNF4A*, a type 2 diabetes–associated gene,[33] members of the

Table 2
Functional enrichment analysis of NAFLD candidate genes associated loci: predicted biological process

Gene Ontology (GO) ID	GO: Biological Process	Gene List
GO:0030730	Sequestering of triglyceride	*ENPP1 IL1B PPARA PPARG TNF*
GO:0006641	Triglyceride metabolic process	*ACSL4 APOC3 APOE DGAT1 DGAT2 LIPC MTTP* ***PNPLA3***
GO:0008610	Lipid biosynthetic process	*ABCB11 ACSL4 APOC3 APOE CFTR DGAT1 DGAT2 FDFT1 LIPC MIF PDGFA PEMT* ***PNPLA3*** *PTGS2 TNF*
GO:0045834	Positive regulation of lipid metabolic process	*ADIPOQ AGTR1 APOE IL1B IRS1 PPARA PPARG PPARGC1A PTGS2 TNF*
GO:0009891	Positive regulation of biosynthetic process	*ADRB2 APOE CLOCK IL1B IL6 IRS1 KLF6 MC4R NR1I2 PDGFA PPARA PPARG PPARGC1A PTGS2 SOD2 STAT3 TCF7L2 TLR4 TNF*
GO:0006869li	Lipid transport	*ACSL4 ADIPOQ APOC3 APOE CFTR FABP2 IL1B LIPC MTTP PPARA PPARG*
GO:0045598	Regulation of fat cell differentiation	*ADIPOQ ENPP1 IL6 PPARG PTGS2 SOD2 TNF*
GO:0009893	Positive regulation of metabolic process	*ADIPOQ ADRB2 ADRB3 AGTR1 APOE CLOCK GCKR GCLC IL1B IL6 IRS1 KLF6 MC4R MIF NR1I2 PDGFA PEMT PPARA PPARG PPARGC1A PTGS2 SOD2 STAT3 TCF7L2 TLR4 TNF*
GO:0044255	Cellular lipid metabolic process	*ACSL4 ADIPOQ ADIPOR1 ADIPOR2 AGTR1 APOC3 APOE CYP2E1 DGAT1 DGAT2 FABP2 FDFT1 IRS1 LIPC MIF MTTP PDGFA PEMT* ***PNPLA3*** *PPARA PPARG PPARGC1A PTGS2 UCP1 UGT1A1*
GO:0050727	Regulation of inflammatory response	*ADIPOQ ADRB2 AGTR1 APOE IL1B IL6 MIF PPARG PTGS2 TLR4 TNF*
Common physiologic processes link the pathophysiology of NAFLD with metabolic syndrome		
GO:0008217	Regulation of blood pressure	*DIPOQ ADRB2 ADRB3 AGTR1 MIF PPARA PPARG PTGS2 SOD2*
GO:0003013	Circulatory system process	*ADIPOQ ADRB2 ADRB3 AGTR1 APOE CFTR GCLC MIF MTHFR PPARA PPARG PTGS2 SOD2*
GO:0008015	Blood circulation	*ADIPOQ ADRB2 ADRB3 AGTR1 APOE CFTR GCLC MIF MTHFR PPARA PPARG PTGS2 SOD2*

uncoupling protein family (ie, *UCP3*), which may play an important role in burning fat, STAT proteins, serpins such as angiotensinogen (*AGT*), and other proinflammatory substances. In fact, the participation of the renin-angiotensin system in the pathophysiology of NAFLD was anticipated by the protection against NAFLD previously reported for losartan, an angiotensin-2 type 1 receptor (AT1R) blocker.[34] Details about the whole set of the top 100 prioritized genes are shown in **Fig. 3**.

Once more, an important topic to highlight is that despite the 2-hit hypothesis still being the leading theory guiding current research on NAFLD, biological evidence

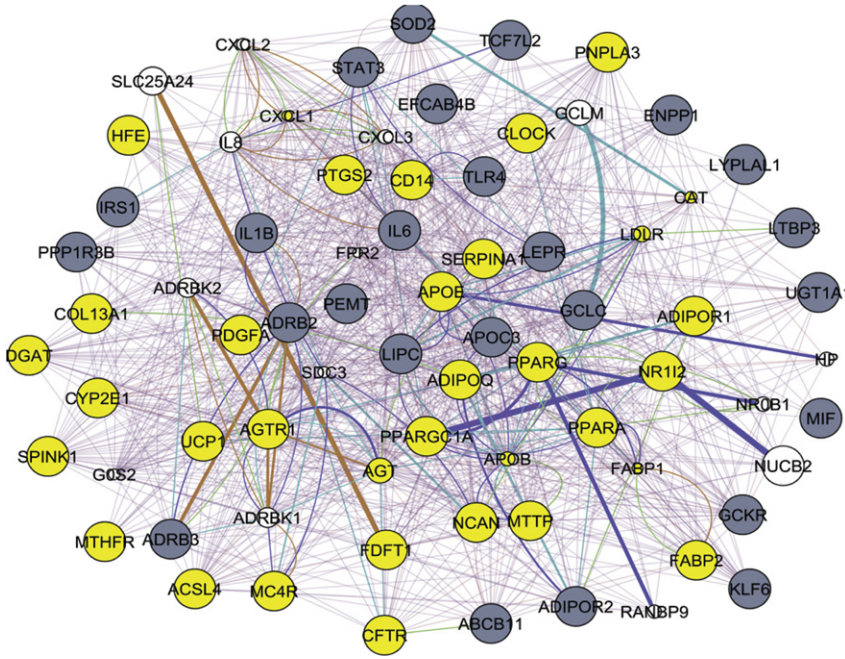

Fig. 2. Functional association analysis of protein and genetic interactions focused on the *PNPLA3* gene. Pathways, coexpression, colocalization, and protein domain similarity were analyzed by the bioinformatic resource GenMANIA. Yellow nodes (genes) are considered to be direct "neighbors" of the *PNPLA3*. The remaining gray circles are the listed genes in **Table 1**. In addition, 20 new predicted genes are incorporated (*white circles*) in the network. Edges interconnecting nodes mean: coexpression, violet (weight: 70.2%); colocalization, brown-bordeaux (11.9%); physical interactions, deep blue (11.1%); pathways, light blue-turquoise (3.4%); shared protein domains, green (0.2%).

strongly suggests that there are other impaired metabolic scenarios that play a key role in the development and progression of NAFLD, for instance, hepatocellular hypoxia, which may lead to the generation of reactive oxygen species, mitochondrial changes, and inflammation along with the associated insulin resistance state.[21]

GENETIC ARCHITECTURE OF NAFLD. CAN GENES FROM DISEASE-ASSOCIATED VARIANTS HELP US TO SELECT THERAPEUTIC TARGETS? TRANSLATING FINDINGS OF NAFLD-RISK GENE VARIANTS INTO THERAPEUTIC TARGET IDENTIFICATION

The tremendous progress on the knowledge of the human genome variation and the explosion of studies exploring the genetic component of a disease offer the unique opportunity to translate this information into the design and identification of promising therapeutic targets. Thus the authors used the available information about gene variants associated with NAFLD risk to address this challenge, and some interesting findings were observed. For example, a remarkable observation concerns drug prediction from functional enrichment analysis. After using the candidate gene list (see **Table 1**) it was found that rosiglitazone, troglitazone, and vitamin E were not only highly predicted, but clustered more than 20 genes out of the 58 reported loci (**Table 3**). Therefore, one might infer that the use of these drugs is strongly supported not only by the

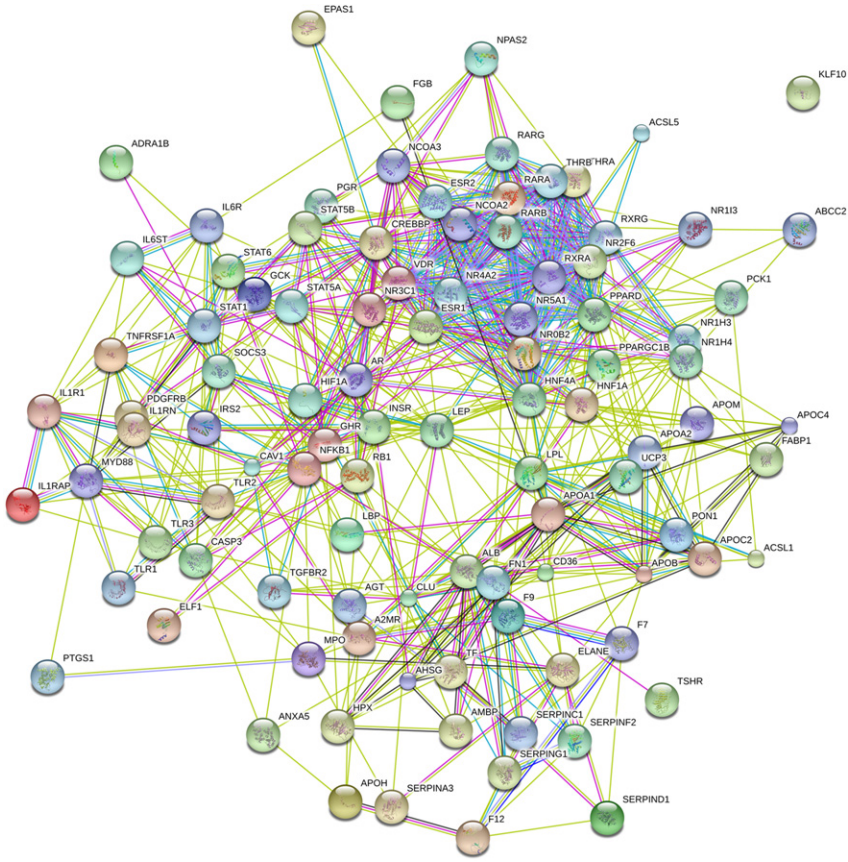

Fig. 3. Best candidates of the whole human genome prioritization for the evaluation of the genetic component of NAFLD. Prioritization was built by the bioinformatic tool ENDEAVOUR, and the figure shows results of the cluster analysis of the first 100 prioritized candidate genes from the whole human genome (23,712 genes) with a significant association with the training set. Analysis was performed using the STRING resource (Search Tool for the Retrieval of Interacting Genes/Proteins), which is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) associations.

clinical phenotype but also by the associated disease gene variants analyzed in an integrative manner. This concept promptly redirects one to the straightforward question: can variations in key genes that encode proteins of disease-associated pathways influence the treatment response? The answer to this question is critical introducing personalized medicine, and some ideas are outlined as follows.

The Web resource STITCH (Search Tool for Interactions of Chemicals; http://stitch.embl.de/), which explores known and predicted interactions of chemicals and proteins, was used to look for interactions of vitamin E. Through this resource, chemicals or drugs are linked to other chemicals and proteins by evidence derived from experiments, databases, and the literature. These associations create a network that can be used to predict and further explore putative mechanisms or putative genes involved in the compound action, among other things. The results of interactions of

Table 3
Functional enrichment analysis of NAFLD candidate genes associated loci: predicted drug

Database ID	Predicted Drug	Gene List
CID000077999	Rosiglitazone	*ABCB11 ACSL4 ADIPOQ ADIPOR1 ADIPOR2 AGTR1 APOC3 APOE CFTR DGAT2 GCKR IL1B IL6 IRS1 MTTP NR1I2* ***PNPLA3*** *PPARA PPARG PPARGC1A PTGS2 SERPINA1 SPINK1 TNF UCP1*
CID000005591	Troglitazone	*ABCB11 ACSL4 ADIPOQ ADIPOR1 ADIPOR2 APOE CYP2E1 ENPP1 IL1B IL6 IRS1 LIPC NR1I2* ***PNPLA3*** *PPARA PPARG PPARGC1A PTGS2 SOD2 STAT3 TNF TNFSF10 UCP1 UGT1A1*
CID000004829	Pioglitazone	*ABCB11 ACSL4 ADIPOQ ADIPOR1 ADIPOR2 AGTR1 APOC3 APOE CYP2E1 IL6 IRS1 LIPC MTTP PPARA PPARG PTGS2 SOD2 STAT3 TCF7L2 TNF TNFSF10 UCP1*
CID000004091	Metformin	*ADIPOQ ADIPOR2 AGTR1 ENPP1 IL6 IRS1 LIPC PPARA PPARG PPARGC1A PTGS2 TCF7L2 TNF*
CID000005056	Resveratrol	*ADIPOQ APOC3 APOE CYP2E1 GCLC IL1B IL6 KLF6 NR1I2 PPARA PPARG PPARGC1A PTGS2 SOD2 STAT3 TNF TNFSF10 UGT1A1*
CID000002116	Vitamin E	*APOC3 APOE CYP2E1 GCLC HFE IL1B IL6 LIPC MTTP NR1I2* ***PNPLA3*** *PPARG PTGS2 SOD2 TNF*
D003474	Curcumin	*ADIPOQ CD14 CFTR CYP2E1 GCLC IL1B IL6 KLF6 MIF NR1I2 PPARG PTGS2 SOD2 STAT3 TCF7L2 TLR4 TNF TNFSF10 UGT1A1*
CID000001106	Stearoyl-coenzyme A	*ADIPOQ ADIPOR2 DGAT2 FDFT1 GCKR LEPR MC4R* ***PNPLA3*** *PPARA PPARG PTGS2 STAT3*
D011794	Quercetin	*APOE CFTR CYP2E1 GCLC IL1B IL6 MIF NR1I2* ***PNPLA3*** *PPARA PPARG PPARGC1A PTGS2 SOD2 STAT3 TCF7L2 TNF TNFSF10 UGT1A1*
CID000000778	Homocysteine	*ADIPOQ APOC3 APOE GCLC IL6 MIF MTHFR PEMT* ***PNPLA3*** *PPARA PPARG SOD2 TNF*
CID000000965	Oleic acid	*ACSL4 ADIPOR1 ADIPOR2 APOC3 APOE DGAT1 DGAT2 FABP2 LIPC MTTP* ***PNPLA3*** *PPARA PPARG PPARGC1A PTGS2 SOD2 UCP1*
CID000000303	Bile acid	*ABCB11 APOC3 APOE CYP2E1 ENPP1 FDFT1 LIPC MTTP NR1I2 PPARA PPARG PPARGC1A PTGS2 UGT1A1*
D007213	Indomethacin	*AGTR1 CYP2E1 GCLC IL1B IL6 PPARA PPARG PTGS2 SERPINA1 SOD2 TCF7L2 TLR4 TNF UGT1A1*
Compounds Associated with Drug-Induced Fatty Liver		
CID000003285	Estrogen	*ABCB11 ADIPOQ ADRB2 ADRB3 AGTR1 APOC3 APOE CFTR CYP2E1 IL1B IL6 IRS1 LIPC NR1I2 PEMT PPARA PPARG PPARGC1A PTGS2 SERPINA1 SOD2 STAT3 TNF UGT1A1*
D011374	Progesterone	*ADIPOQ ADIPOR2 ADRB2 ADRB3 CFTR CYP2E1 ENPP1 FDFT1 IL1B IL6 IRS1 NR1I2 PDGFA PPARG PPARGC1A PTGS2 SOD2 TLR4 TNF TNFSF10 UGT1A1*
CID000003003	Dexamethasone	*ABCB11 ADIPOQ ADIPOR2 ADRB2 ADRB3 AGTR1 APOC3 APOE CD14 CYP2E1 ENPP1 GCLC IL1B IL6 IRS1 MC4R MIF NR1I2* ***PNPLA3*** *PPARA PPARG PPARGC1A PTGS2 SERPINA1 SOD2 STAT3 TCF7L2 TNF TNFSF10 UGT1A1*

Prediction of related drugs was performed based on annotations from the following resources: CTD (Comparative Toxicogenomics Database), CTD Marker, CTD Therapeutic, Drug Bank, and STITCH (Search Tool for Interactions of Chemicals).

vitamin E are shown in **Fig. 4**. Vitamin E shows a chemical-protein interaction with CYP4F2 (cytochrome P450, family 4, subfamily F, polypeptide 2) which, in liver microsomes, is involved in a nicotinamide adenine dinucleotide phosphate–dependent electron transport pathway. Moreover, CYP4F2 oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics, and starts the process of inactivating and degrading leukotriene B_4, a potent mediator of inflammation. In addition, vitamin E also shows chemical-protein interactions with PLA2G1B (phospholipase A_2), group IB (pancreas) that catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-*sn*-phosphoglycerides releasing arachidonic acid, whose importance was aforementioned; APOB (apolipoprotein B), a major protein constituent of chylomicrons; PRKCA (protein kinase C, α), a calcium-activated, phospholipid-dependent, serine- and threonine-specific enzyme; ALOX5 (arachidonate 5-lipoxygenase Z), which catalyzes the first step in leukotriene biosynthesis; SOD2 (superoxide dismutase 2, mitochondrial); and GSR (glutathione reductase), which maintains high levels of reduced glutathione in the cytosol (see **Fig. 4**).

Perhaps this information could be used as a prognostic and predictive tool to guide therapeutic decisions. Looking at genetic variants of *CYP4F2*, perhaps one may better predict patients' treatment response and distinguish between responders and nonresponders to therapy. However, this is an open question that needs to be answered in future research and future clinical trials.

In addition, the authors inquired as to which of the highly predicted drugs or chemical components incorporated *PNPLA3* in the functional analysis (see **Table 3**). *PNPLA3* was found to be associated with some of the already proven pharmacologic agents, with reasonable efficacy on improving fatty liver, such as rosiglitazone and vitamin E,[35] but also with some other compounds that have never been tested in

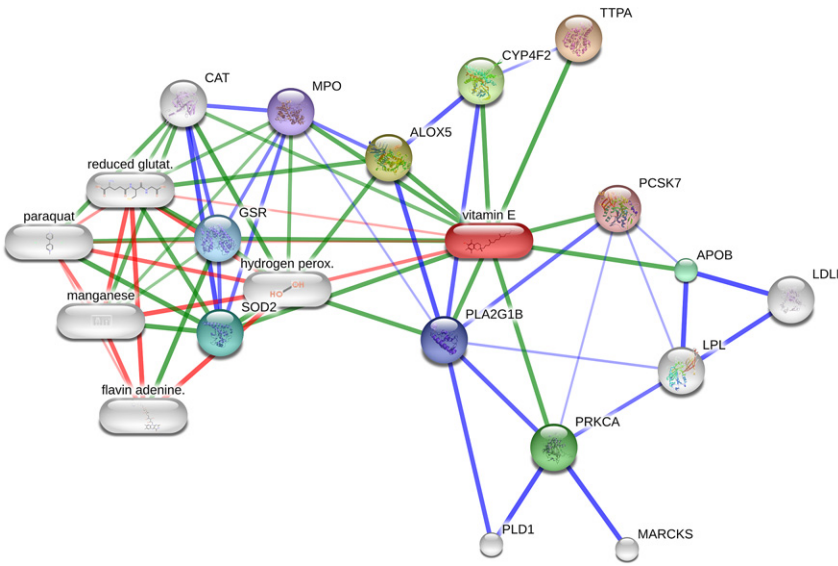

Fig. 4. In silico prediction of interactions between drugs and target proteins: Network around Vitamin E. Network prediction was performed by the STITCH resource. Protein-protein interactions are shown in blue, chemical-protein interactions in green, and interactions between chemicals in red. Chemicals are represented as pill-shaped nodes, whereas proteins are shown as spheres.

humans such as quercetin, a flavonoid widely distributed in nature. Despite this information, more thorough research is needed because some evidence in previous experimental works showed that flavonoids operating on the redox state can improve diet-induced fatty liver,[36] and quercetin decreases de novo fatty acid and TAG synthesis in rat hepatocytes.[37] Some other compounds, such as homocysteine and oleic acid, were also highly predicted in association with *PNPLA3* (see **Table 3**).

Not surprisingly, compounds that are well known to be associated with drug-induced fatty liver, such as corticosteroids and sexual hormones, were also highly predicted in association with a great majority of the target genes, including *PNPLA3* (see **Table 3**), suggesting that the pathogenesis of abnormal liver triglyceride accumulation involves common disease pathways.

In addition, the coactivator *PPARGC1A*, with its transcriptional activity modulated by epigenetic factors, and which seems to participate in the pathogenesis of insulin resistance and NAFLD,[38] was highly predicted by most of the drug compounds (see **Table 3**).

Finally, the transcription factor *STAT3*, previously reported as being involved in the genetic susceptibility of NAFLD progression and severity,[2] was also predicted for several compounds, including resveratrol, a well-known compound explored in animals but scarcely in humans, with beneficial effects on the various aspects of body homeostasis.[39]

GENES AND NAFLD: MODERN MEDICINE AND FUTURE CHALLENGES

The understanding of the genetic risk of complex diseases opens new opportunities and challenges and more practical questions. For example, can we rapidly translate the knowledge about NAFLD genetic susceptibility into more individualized decision making and personalized medicine? If so, can we improve the diagnostic and therapeutic strategies?

The ideal situation would be the design of tools based on the patient's genetic profile that would allow improved diagnosis and tailored treatment without invasive procedures and with minimal adverse events. Perhaps the clearest attempt to combine genetic data and clinical features was performed by Kotronen and coworkers.[40] Their study evaluated the performance of predicting NAFLD by combining routine clinical and laboratory data and the rs738409 genotypes. Surprisingly, however, the addition of the rs738409 patient's information into the disease modeling to create a score only improved the accuracy of the prediction by less than 1%.

An interesting and attractive use of the genetic data as mentioned previously is the idea to replace invasive procedures such as liver biopsy by risk alleles. Unfortunately, in this regard the incorporation of genetic information is not yet very promising, as traditional and noncanonical risk factors, when combined properly, have a good predictive power.[41] Hence, we would have to rely on genetic variants with stronger risk effects to observe better results in terms of disease prediction, in comparison with the current serologic biomarkers or even the more simple algorithms.[42]

Assuming that both environmental and genetic influences are likely involved in the NAFLD treatment response to a given drug, can we expect that all patients will equally respond to the therapy? Certainly we cannot. In this framework, gene variants with a proven effect on the biology and natural history of the disease, such as rs738409, could be included in future tailored therapeutic regimens to better predict, for instance, drug response. In this context, one may speculate that carriers of the G allele of the rs738409 variant might show less improvement in histologic outcomes compared with homozygous for the C allele. Thus, it would

be interesting to stratify the treatment response and, in particular, adjust drug doses to the patient genotype. Conversely, as already proved for a myriad of drug targets, namely, angiotensin I–converting enzyme (peptidyl-dipeptidase A or ACE) or AT1R, carriers of the G allele of the rs738409 variant might show a better response for a drug designed for binding to the PNPLA3 protein. In fact, in vitro assays using recombinant PNPLA3 showed that the wild-type enzyme hydrolyzes emulsified triglycerides, and the 148M substitution abolishes this activity.[43] Structural protein modeling also predicted that this substitution may restrict access of substrate to the catalytic serine at residue 47 by enhancing the distance between the proximal Asp^{166} and Cys^{47}, as shown in **Fig. 5**. More studies using crystallographic analysis should be done to confirm these assumptions and test putatively active drugs.

THE POTENTIAL ADVANCE GENERATED BY NEXT-GENERATION SEQUENCING TECHNIQUES

Great promise has been raised by the emergence of high-throughput next-generation sequencing technologies, initially applied, but not restricted, to the discovery of rare variants causing Mendelian diseases not only in the exome but also in the

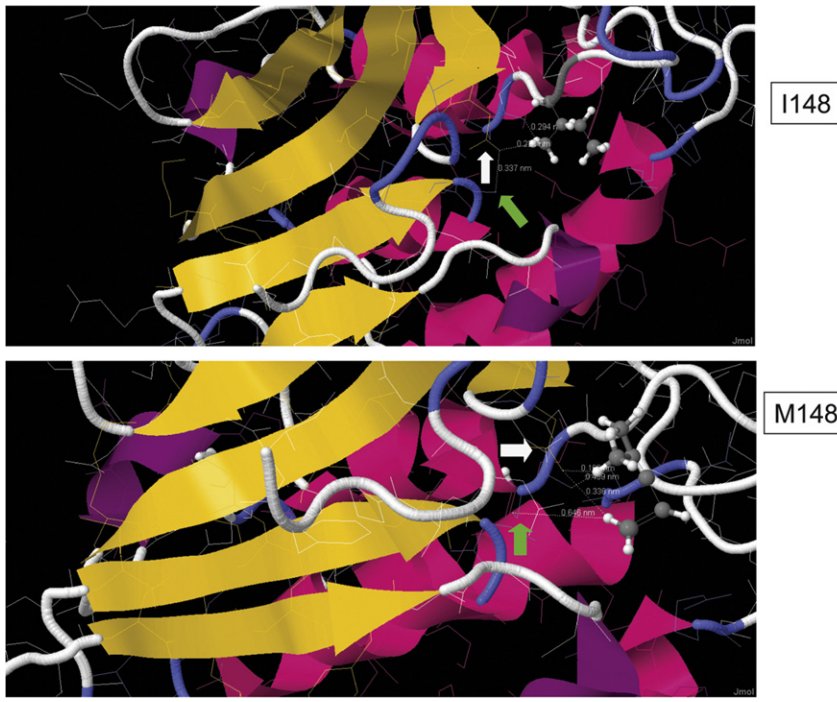

Fig. 5. The use of genetic information to optimize drug therapy: Structure protein modeling of *PNPLA3* rs738409. Structural protein prediction shows that the I148M substitution may restrict access of substrate to the catalytic serine at residue 47 by enhancing the distance between the proximal Asp^{166} and Cys^{47}. Green arrows show Ser^{47} and white arrows show Asp^{166}. Prediction was built using the SWISS-MODEL resource at http://swissmodel.expasy.org/, and visualization of chemical structure in 3D was prepared in Jmol at http://www.jmol.org/.

whole genome, as the techniques have become more available and cheaper. In fact, the 1000 Genome Project (http://browser.1000genomes.org) will contribute powerful information to the better understanding of human genetic variation. In addition, exome sequencing is being adapted to explore the contribution of rare variants (present in <1% of the population) to the "missing" heritability of complex diseases. It is commonly stated that association studies "discover disease genes," a misinterpretation that is taken too often. The fact is that associated markers map to block linked SNPs that may extend more than 100 kb and may contain several genes, and there is no guarantee of the functional involvement of the variant in the genetic effect. Fine mapping to the functional variant requires the identification of all candidate polymorphisms, obtaining the sequence of the entire region for the discovery of all the variants to explain the association in a large number of cases and controls. Moreover, as discussed recently,[44] next-generation sequencing is faced with problems in the selection of the best design to obtain good statistical power, the method of filtering neutral variants, and subsequent identification of true disease-associated polymorphisms and the development of efficient bioinformatics algorithms for the management of this huge amount of information.

Haplotype analysis of *PNPLA3* shows that the rs738409 is in moderate linkage disequilibrium (LD) (r^2 = 0.65) with the other variants, including the rs6006460 and rs2294918. Thus, this scenario precludes any imputation across the *PNPLA3* locus centered on the rs738409, and suggests that the I148M variant might be the causal variant in the susceptibility of fatty liver. Nevertheless, annotation of nearby SNPs in LD (proxies) with the rs738409 based on HapMap data and nearby loci shows other loci, such as SAMM50, which is a component of the sorting and assembly machinery (SAM) complex of the outer mitochondrial membrane. Of interest is that a nonsynonymous SNP (rs3761472) in *SAMM50* was associated with elevated liver enzymes,[45] suggesting that SAMM50 could be a good candidate for follow-up studies.

BEYOND DNA SEQUENCE VARIATION: THE ROLE OF EPIGENETICS IN NAFLD

Other genetic factors beyond DNA sequence variation play a critical role in the etiology of NAFLD and could potentially explain the interaction between the disease, the environmental influences, and other phenotypes. Among these factors, epigenetic modifications are the best candidates.

Epigenetics is important in understanding the pathophysiology of NAFLD because, in addition to biological plausibility, epigenetic modifications regulate chromosome organization and gene transcription by definition and are, in turn, highly controlled by environmental stimuli, including nutritional status; are highly dynamic; and can occur de novo but are potentially transmitted to the next generation.

In this context, the authors hypothesize that NAFLD is intimately associated with the status of insulin resistance and that both conditions are strongly linked by local epigenetic modifications occurring in the liver tissue before or after fat transformation. The authors observed that methylation levels of the promoter of the transcriptional coactivator, peroxisome proliferator-activated receptor γ coactivator 1α (*PPARGC1A*), in the liver tissue of patients with NAFLD correlate with HOMA-IR (homeostasis model assessment–estimated insulin resistance) and plasma fasting insulin levels.[38] Accordingly, it was also observed that liver abundance of *PPARGC1A* mRNA is inversely correlated with the methylation levels of *PPARGC1A* promoter, and also with the status of peripheral insulin resistance, suggesting that methylation of certain points in the gene promoter efficiently repressed its transcriptional activity.[38] Finally,

mitochondrial biogenesis is reduced in the liver of NAFLD patients and is associated with peripheral insulin resistance and *PPARGC1A* promoter methylation status. These changes can be observed in the umbilical cord of small- and large-for-gestational-age newborns,[46,47] which may explain the fetal origin of adult disease by metabolic programming, as originally proposed by Barker and recently reviewed by Dyer and Rosenfeld.[48]

To conclude, after the description by Romeo et al[6] of the association of the I148M variant with fatty liver, we replicated the findings for the first time and extended the association to disease severity in biopsy-proven NAFLD patients,[49] which was replicated in a series of different studies.[7] These findings, as shown here, may open new opportunities for the understanding of the disease pathophysiology and the design of new therapies.

REFERENCES

1. Hernaez R. Genetic factors associated with the presence and progression of nonalcoholic fatty liver disease: a narrative review. Gastroenterol Hepatol 2012; 35(1):32–41.
2. Sookoian S, Castano G, Gianotti TF, et al. Genetic variants in STAT3 are associated with nonalcoholic fatty liver disease. Cytokine 2008;44(1):201–6.
3. Sookoian S, Castano GO, Burgueno AL, et al. The nuclear receptor PXR gene variants are associated with liver injury in nonalcoholic fatty liver disease. Pharmacogenet Genomics 2010;20(1):1–8.
4. Carulli L, Canedi I, Rondinella S, et al. Genetic polymorphisms in non-alcoholic fatty liver disease: interleukin-6-174G/C polymorphism is associated with non-alcoholic steatohepatitis. Dig Liver Dis 2009;41(11):823–8.
5. Namikawa C, Shu-Ping Z, Vyselaar JR, et al. Polymorphisms of microsomal triglyceride transfer protein gene and manganese superoxide dismutase gene in non-alcoholic steatohepatitis. J Hepatol 2004;40(5):781–6.
6. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40(12):1461–5.
7. Sookoian S, Pirola CJ. Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. Hepatology 2011;53(6):1883–94.
8. Bambha K, Belt P, Abraham M, et al. Ethnicity and nonalcoholic fatty liver disease. Hepatology 2012;55(3):769–80.
9. Huang Y, Cohen JC, Hobbs HH. Expression and characterization of a PNPLA3 protein isoform (I148M) associated with nonalcoholic fatty liver disease. J Biol Chem 2011;286(43):37085–93.
10. Prentki M, Madiraju SR. Glycerolipid metabolism and signaling in health and disease. Endocr Rev 2008;29(6):647–76.
11. Zechner R, Strauss JG, Haemmerle G, et al. Lipolysis: pathway under construction. Curr Opin Lipidol 2005;16(3):333–40.
12. Kienesberger PC, Oberer M, Lass A, et al. Mammalian patatin domain containing proteins: a family with diverse lipolytic activities involved in multiple biological functions. J Lipid Res 2009;50(Suppl):S63–8.
13. Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 2010;38(Web Server issue):W214–20.
14. Puri P, Wiest MM, Cheung O, et al. The plasma lipidomic signature of nonalcoholic steatohepatitis. Hepatology 2009;50(6):1827–38.

15. Basantani MK, Sitnick MT, Cai L, et al. Pnpla3/Adiponutrin deficiency in mice does not contribute to fatty liver disease or metabolic syndrome. J Lipid Res 2011;52(2):318–29.
16. Chen W, Chang B, Li L, et al. Patatin-like phospholipase domain-containing 3/adiponutrin deficiency in mice is not associated with fatty liver disease. Hepatology 2010;52(3):1134–42.
17. Yang LY, Kuksis A, Myher JJ, et al. Origin of triacylglycerol moiety of plasma very low density lipoproteins in the rat: structural studies. J Lipid Res 1995;36(1):125–36.
18. Jenkins CM, Mancuso DJ, Yan W, et al. Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities. J Biol Chem 2004;279(47):48968–75.
19. Turban S, Hajduch E. Protein kinase C isoforms: mediators of reactive lipid metabolites in the development of insulin resistance. FEBS Lett 2011;585(2):269–74.
20. Li ZZ, Berk M, McIntyre TM, et al. Hepatic lipid partitioning and liver damage in nonalcoholic fatty liver disease: role of stearoyl-CoA desaturase. J Biol Chem 2009;284(9):5637–44.
21. Carabelli J, Burgueno AL, Rosselli MS, et al. High fat diet-induced liver steatosis promotes an increase in liver mitochondrial biogenesis in response to hypoxia. J Cell Mol Med 2011;15(6):1329–38.
22. Sookoian S, Gianotti TF, Rosselli MS, et al. Liver transcriptional profile of atherosclerosis-related genes in human nonalcoholic fatty liver disease. Atherosclerosis 2011;218(2):378–85.
23. Baulande S, Lasnier F, Lucas M, et al. Adiponutrin, a transmembrane protein corresponding to a novel dietary- and obesity-linked mRNA specifically expressed in the adipose lineage. J Biol Chem 2001;276(36):33336–44.
24. D'Eon TM, Souza SC, Aronovitz M, et al. Estrogen regulation of adiposity and fuel partitioning. Evidence of genomic and non-genomic regulation of lipogenic and oxidative pathways. J Biol Chem 2005;280(43):35983–91.
25. Rosselli MS, Burgueno AL, Pirola CJ, et al. Cyclooxygenase inhibition upregulates liver carnitine palmitoyltransferase 1A expression and improves fatty liver. Hepatology 2011;53(6):2143–4.
26. Aerts S, Lambrechts D, Maity S, et al. Gene prioritization through genomic data fusion. Nat Biotechnol 2006;24(5):537–44.
27. Sookoian S, Gianotti TF, Schuman M, et al. Gene prioritization based on biological plausibility over genome wide association studies renders new loci associated with type 2 diabetes. Genet Med 2009;11(5):338–43.
28. Lefebvre P, Cariou B, Lien F, et al. Role of bile acids and bile acid receptors in metabolic regulation. Physiol Rev 2009;89(1):147–91.
29. Knop FK. Bile-induced secretion of glucagon-like peptide-1: pathophysiological implications in type 2 diabetes? Am J Physiol Endocrinol Metab 2010;299(1):E10–3.
30. Ding X, Saxena NK, Lin S, et al. Exendin-4, a glucagon-like protein-1 (GLP-1) receptor agonist, reverses hepatic steatosis in ob/ob mice. Hepatology 2006; 43(1):173–81.
31. Byrne CD. Hypoxia and non-alcoholic fatty liver disease. Clin Sci (Lond) 2010; 118(6):397–400.
32. Nath B, Szabo G. Hypoxia and hypoxia inducible factors: diverse roles in liver diseases. Hepatology 2012;55(2):622–33.
33. Sookoian S, Gemma C, Pirola CJ. Influence of hepatocyte nuclear factor 4alpha (HNF4alpha) gene variants on the risk of type 2 diabetes: a meta-analysis in 49,577 individuals. Mol Genet Metab 2010;99(1):80–9.

34. Rosselli MS, Burgueno AL, Carabelli J, et al. Losartan reduces liver expression of plasminogen activator inhibitor-1 (PAI-1) in a high fat-induced rat nonalcoholic fatty liver disease model. Atherosclerosis 2009;206(1):119–26.
35. Sanyal AJ, Chalasani N, Kowdley KV, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. N Engl J Med 2010;362(18):1675–85.
36. Rapavi E, Kocsis I, Feher E, et al. The effect of citrus flavonoids on the redox state of alimentary-induced fatty liver in rats. Nat Prod Res 2007;21(3):274–81.
37. Gnoni GV, Paglialonga G, Siculella L. Quercetin inhibits fatty acid and triacylglycerol synthesis in rat-liver cells. Eur J Clin Invest 2009;39(9):761–8.
38. Sookoian S, Rosselli MS, Gemma C, et al. Epigenetic regulation of insulin resistance in nonalcoholic fatty liver disease: impact of liver methylation of the peroxisome proliferator-activated receptor gamma coactivator 1alpha promoter. Hepatology 2010;52(6):1992–2000.
39. Smoliga JM, Baur JA, Hausenblas HA. Resveratrol and health—a comprehensive review of human clinical trials. Mol Nutr Food Res 2011;55(8):1129–41.
40. Kotronen A, Peltonen M, Hakkarainen A, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. Gastroenterology 2009;137(3):865–72.
41. Adams LA, Feldstein AE. Non-invasive diagnosis of nonalcoholic fatty liver and nonalcoholic steatohepatitis. J Dig Dis 2011;12(1):10–6.
42. Sookoian S, Castano G, Burgueno AL, et al. A diagnostic model to differentiate simple steatosis from nonalcoholic steatohepatitis based on the likelihood ratio form of Bayes theorem. Clin Biochem 2009;42(7–8):624–9.
43. He S, McPhaul C, Li JZ, et al. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. J Biol Chem 2010;285(9):6706–15.
44. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 2011;12(11):745–55.
45. Yuan X, Waterworth D, Perry JR, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet 2008;83(4):520–8.
46. Gemma C, Sookoian S, Alvarinas J, et al. Mitochondrial DNA depletion in small- and large-for-gestational-age newborns. Obesity (Silver Spring) 2006;14(12):2193–9.
47. Gemma C, Sookoian S, Alvarinas J, et al. Maternal pregestational BMI is associated with methylation of the PPARGC1A promoter in newborns. Obesity (Silver Spring) 2009;17(5):1032–9.
48. Dyer JS, Rosenfeld CR. Metabolic imprinting by prenatal, perinatal, and postnatal overnutrition: a review. Semin Reprod Med 2011;29(3):266–76.
49. Sookoian S, Castaño GO, Burgueño AL, et al. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res 2009;50(10):2111–6.