Dear Author

Here are the proofs of your article.

- You can submit your corrections **online,** via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For **fax** submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style.
- Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL:

`http://dx.doi.org/10.1007/s11901-014-0222-x`

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information, go to: http://www.link.springer.com.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us, if you would like to have these documents returned.

Metadata of the article that will be visualized in OnlineFirst

1	Article Title	**Genetic and Epigenetic Associations with NAFLD: Focus on Clinical Decision Making and Novel Concepts in Disease Pathogenesis**	
2	Article Sub- Title		
3	Article Copyright - Year	**Springer Science+Business Media New York 2014 (This will be the copyright line in the final PDF)**	
4	Journal Name	Current Hepatology Reports	
5	Corresponding Author	Family Name	**Sookoian**
6		Particle	
7		Given Name	**Silvia**
8		Suffix	
9		Organization	Instituto de Investigaciones Médicas A. Lanari-CONICET
10		Division	
11		Address	Combatiente de Malvinas 3150, Buenos Aires 1427, Argentina
12		Organization	University of Buenos Aires – National Scientific and Technical Research Council (CONICET)
13		Division	Department of Clinical and Molecular Hepatology, Institute of Medical Research A Lanari-IDIM
14		Address	Ciudad Autónoma de Buenos Aires, Argentina
15		e-mail	sookoian.silvia@lanari.fmed.uba.ar
16	Corresponding Author	Family Name	**Pirola**
17		Particle	
18		Given Name	**Carlos J.**
19		Suffix	
20		Organization	Instituto de Investigaciones Médicas A. Lanari-CONICET
21		Division	
22		Address	Combatiente de Malvinas 3150, Buenos Aires 1427, Argentina
23		Organization	University of Buenos Aires – National Scientific and Technical Research Council (CONICET)

24		Division	Department of Molecular Genetics and Biology of Complex Diseases, Institute of Medical Research A Lanari-IDIM
25		Address	Ciudad Autónoma de Buenos Aires, Argentina
26		e-mail	pirola.carlos@lanari.fmed.uba.ar

27		Received	
28	Schedule	Revised	
29		Accepted	

30	Abstract	Nonalcoholic fatty liver disease (NAFLD) is a complex liver disease with worldwide prevalence. Its development involves a myriad of factors, including genetic susceptibility and environmental insults. In this review, we summarize new findings about current genome-wide association studies on NAFLD. In addition, we used a strategy of functional enrichment analysis to integrate all the newly discovered loci into common biological pathways and to explore their role in the pathogenesis of NAFLD. Controversies on the application of genetic testing to predict disease severity are discussed and specifically the role of rs738409 in clinical decision making. Finally, we highlighted significant trends and developments in epigenetic changes and microRNAs associated with disease progression.

31	Keywords separated by ' - '	NAFLD - NASH - Genetics - PNPLA3 - Personalized medicine - GWAS - Risk prediction - Epigenetics - DNA methylation

32	Foot note information	

Curr Hepatology Rep
DOI 10.1007/s11901-014-0222-x

FATTY LIVER DISEASE (SA HARRISON AND J GEORGE, SECTION EDITORS)

Genetic and Epigenetic Associations with NAFLD: Focus on Clinical Decision Making and Novel Concepts in Disease Pathogenesis

Silvia Sookoian · Carlos J. Pirola

Abstract Nonalcoholic fatty liver disease (NAFLD) is a complex liver disease with worldwide prevalence. Its development involves a myriad of factors, including genetic susceptibility and environmental insults. In this review, we summarize new findings about current genome-wide association studies on NAFLD. In addition, we used a strategy of functional enrichment analysis to integrate all the newly discovered loci into common biological pathways and to explore their role in the pathogenesis of NAFLD. Controversies on the application of genetic testing to predict disease severity are discussed and specifically the role of rs738409 in clinical decision making. Finally, we highlighted significant trends and developments in epigenetic changes and microRNAs associated with disease progression.

Keywords NAFLD · NASH · Genetics · PNPLA3 · Personalized medicine · GWAS · Risk prediction · Epigenetics · DNA methylation

S. Sookoian
Department of Clinical and Molecular Hepatology, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires – National Scientific and Technical Research Council (CONICET), Ciudad Autónoma de Buenos Aires, Argentina

C. J. Pirola
Department of Molecular Genetics and Biology of Complex Diseases, Institute of Medical Research A Lanari-IDIM, University of Buenos Aires – National Scientific and Technical Research Council (CONICET), Ciudad Autónoma de Buenos Aires, Argentina

S. Sookoian (✉) · C. J. Pirola (✉)
Instituto de Investigaciones Médicas A. Lanari-CONICET, Combatiente de Malvinas 3150, Buenos Aires 1427, Argentina
e-mail: sookoian.silvia@lanari.fmed.uba.ar
e-mail: pirola.carlos@lanari.fmed.uba.ar

Abbreviations

ALT	Alanine-aminotransferase
CT	Computed tomography
GWAS	Genome-wide association study
IR	Insulin resistance
MetSyn	Metabolic syndrome
miRNA	microRNA
NAFLD	Nonalcoholic fatty liver disease
NASH	Nonalcoholic steatohepatitis
PNPLA3	Patatin-like phospholipase domain containing 3
PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1α
SNP	Single nucleotide polymorphism

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a liver disease that is prevalent worldwide. It develops from a complex process that involves a myriad of factors, including individual genetic susceptibility and particular environmental insults. Until a few years ago, our knowledge about the genetic components of NAFLD and nonalcoholic steatohepatitis (NASH), the more severe clinical form of NAFLD, was based on results from candidate gene association studies that identified several loci associated with disease susceptibility and progression [1]. Although all of these studies were inspired by biological plausibility, only a few of them were replicated. A remarkable breakthrough in our understanding of the genetic susceptibility to NAFLD was however provided by findings from the first genome-wide association (GWAS) study on NAFLD done by the Dallas Heart Study in 2008 [2].

Environmental factors, such as physical activity [3, 4] and diet intervention [4–6], play an important role in the development of NAFLD. Interestingly, evidence from human studies have provided clues about how gene–environment interactions modulated by epigenetic mechanisms impact not only on the pathogenesis of NAFLD but also the modulation of metabolic syndrome (MetSyn)-related phenotypes, including insulin resistance (IR) [7••, 8].

Finally, although by definition NAFLD is characterized by abnormal liver fat accumulation in the absence of significant alcohol consumption and other causes of secondary hepatic steatosis [9], evidence from clinico-epidemiologic [10], histologic [11], and even in silico systems biology of the disease [12] suggests that NAFLD and alcoholic liver disease (ALD) share many disease determinants, including the same underlying genetic risk [13].

Hence, this review summarizes new findings about GWAS on NAFLD and significant trends and developments on the epigenetic component of the disease. In addition, controversies about disease pathogenesis and management derived from recent discoveries on gene variants are also discussed.

GWAS on NAFLD and Our Knowledge About Disease Pathogenesis

The first GWAS on NAFLD was a genome-wide survey of non-synonymous sequence variations encompassing 9229 single nucleotide polymorphisms (SNPs) in a multiethnic population-based study [2]. The authors uncovered a significant association of the patatin-like phospholipase domain containing 3 (*PNPLA3*, also known as adiponutrin) rs738409 C/G variant, encoding an amino acid substitution (I148M) with liver triglyceride accumulation [2]. This association remained significant even after adjusting for common metabolic confounders such as obesity, diabetes status and related risk factors of disease, such as ethanol use.

Soon thereafter, this finding was replicated, and the *PNPLA3*-148 M allele was significantly associated with disease severity [14]. rs738409 is widely acknowledged as the "NASH gene" because the association is largely replicated around the world not only in adults but also in children [15]. Of note, the risk effect of rs738409 on developing fatty liver is perhaps one of the strongest ever worldwide-replicated effect for a common variant modifying the genetic susceptibility to a complex disease (5.3 % of the total variance) [14]. In addition, homozygous GG carriers have a 3.24-fold greater risk of higher necroinflammatory scores and a 3.2-fold greater risk of developing fibrosis when compared with homozygous CC [15]. Interestingly, the genetic models do not seem to be similar for liver fat and disease severity, and, at least for liver fat deposition, the effect of the variant seems to be greater in women than in men [15].

The use of a GWAS strategy in the search for the genetic basis of NAFLD was followed by other reports that included different populations, study designs, sample sizes, and approaches for the characterization of the main liver phenotype. For example, studies have been undertaken of female adults with NAFLD diagnosed by liver biopsy [16], of the heritability of hepatic steatosis at the population level with computed tomography (CT) [17], a combined approach of CT and alanine-aminotransferase (ALT) levels as a surrogate of disease severity [18], and exploration of the genetic risk in Asian-descent patients [19, 20].

It is important to highlight that the coverage of SNPs by the above-mentioned GWA studies was not uniform in terms of the explored variants. In addition, it varied from a GWAS analysis of 12,138 non-synonymous sequence variations from dbSNP and the Perlegen SNP database [2] to commercial platforms, such as HumanCNV370-Quadv3 BeadChip (coverage: 373,397 SNPs) [16] or Illumina Human 610-Quad BeadChip (coverage: 484,751 SNPs), meta-analysis and GWAS association data of large consortiums that used the Affymetrix 6.0 or Illumina platform [17], and imputed SNPs [18].

Finally, the GWAS strategy has also been used to explore the genetic locus that influences liver enzyme levels in the population, including ALT [21, 22]. A summary of the latest GWAS on NAFLD and ALT levels is depicted in Fig. 1. Surprisingly, the most significant findings are on chromosome 22 at *PNPLA3* loci, and rs738409 is still the most consistently replicated variant associated with fatty liver, disease severity, and associated traits, such as ALT levels. Likewise, rs738409 is consistently associated with NAFLD across different populations [15, 23, 24].

Hence, a number of questions emerge from these results. For instance, we may wonder whether these findings are an indication that the genetic risk of NAFLD is so far explained by a single common variant with a minor allele risk frequency of ~30 %. If so, do the findings fit the concept that NAFLD, like many other common complex diseases, including type 2 diabetes or obesity, is a complex trait influenced by the effect of multiple gene variants. The answer is a definitive yes and that, although the effect size of the variant is one of the biggest ever observed for a common SNP, a significant proportion of the heredity of the trait is missing. The question remains whether carriers should be closely monitored for serious complications of NASH, such as hepatocellular carcinoma [25].

Alternatively, we might wonder whether or not rare variants have a place in the genetic predisposition to NAFLD. In this sense, NAFLD, at least up to now, differs from other complex diseases in that no truly monogenic forms (patients with rare genetic variants with penetrance high enough to explain the phenotype) have been described. Unfortunately, there are no data about variants with a minor allele frequency of less than 1 % influencing the susceptibility to NAFLD.

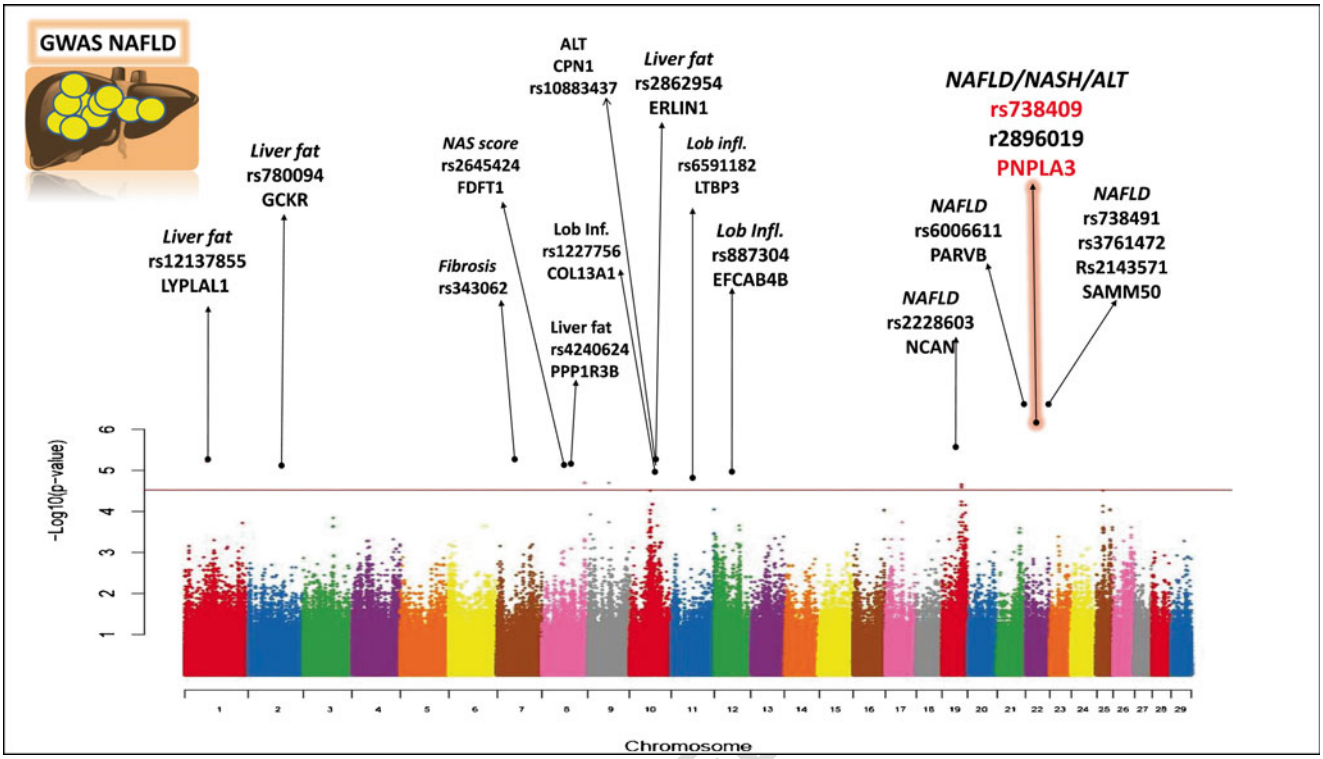

Fig. 1 GWAS on NAFLD: Summary representation of variants significantly associated with NAFLD, NASH, and plasma levels of alanine-aminotransferase (ALT). The illustration resembles a Manhattan plot, with the x-axis denoting the genomic coordinates and chromosome localization of significantly associated SNPs and the y-axis representing the p-value for the association. LYPLAL1 (lysophospholipase-like 1), GCKR (glucokinase (hexokinase 4) regulator), COL13A1 (collagen, type XIII, alpha 1), PPP1R3B (protein phosphatase 1, regulatory subunit 3B), ERLIN1 (ER lipid raft-associated 1), EFCAB4B (EF-hand calcium-binding domain 4B), NCAN (neurocan), PARVB (parvin, beta), PNPLA3 (patatin-like phospholipase domain containing 3), SAMM50 (sorting and assembly machinery component 50 homologue (*S. cerevisiae*), LTBP3 (latent transforming growth factor beta binding protein 3), FDFT1 (farnesyl-diphosphate farnesyltransferase 1), CPN1 (carboxypeptidase N, polypeptide 1)

However, a recent study that did whole-exome sequencing of five loci associated with NAFLD on a small sample of patients with extreme obesity (BMI >50) and NAFLD-related cirrhosis showed that one of four patients was compound heterozygous for putatively rare damaging mutations in *PNPLA3* [26]. Conversely, Cefalu et al. used exome sequencing to discover a novel nonsense mutation in exon 26 of APOB (p.K2240X) responsible for a low cholesterol and fatty liver in a large kindred with familial hypobetalipoproteinemia in which fatty liver is a common feature. This mutation may also be responsible for cirrhosis and liver cancer in this family [27].

Another important question that remains under investigation is the unknown role of non-*PNPLA3*-NAFLD-GWAS-associated variants in the modulation of disease. One may wonder whether or not they have a biological connection either with each other or with the disease. To answer this question, we used a strategy for functional enrichment analysis based on an algorithm that weighs among other options, gene ontology and the underlying biological process to predict in silico a network among the input genes/proteins (GeneMANIA [28]. This strategy not only predicts information about co-expression, physical protein and genetic interaction, co-localization, and common pathways among input genes/proteins, but also extends the list to functionally similar genes. The analysis shows that seven new predicted genes/protein, including GYS1 (glycogen synthase 1, muscle), GYS2 (glycogen synthase 2, liver), PHKG1 (phosphorylase kinase gamma 1, muscle), PHKG2 (phosphorylase kinase gamma 2, testis), PHKB (phosphorylase kinase beta), PHKA1 (phosphorylase kinase alpha 1, muscle), and PHKA2 (phosphorylase kinase alpha 2, liver) were significantly associated with biological pathways that included the glycogen and glucan metabolic process (a series of chemical reactions involving glucans, polysaccharides consisting only of glucose residues) and polysaccharide and carbohydrate metabolism (Table 1).

Table 1 provides detailed information about the interrelated biological functions of the NAFLD-GWAS-associated genes. Surprisingly, *PNPLA3* does not show either genetic interaction, biological pathways, or a shared protein domain with any of the input or newly predicted genes. Indeed, the family of patatin-like phospholipases consists of glycoproteins that account for up to 40 % of the total soluble protein in potato tubers [29]. Remarkably, besides the phospholipase activity,

t1.1 **Table 1** Functional association analysis of protein and genetic interactions focused on genes associated with NAFLD in current GWAS studies

t1.2	Function	FDR	Coverage
t1.3	Glycogen metabolic process	3.01E-9	7 / 43
t1.4	Glucan metabolic process	3.01E-9	7 / 44
t1.5	Cellular glucan metabolic process	3.01E-9	7 / 44
t1.6	Cellular polysaccharide metabolic process	8.99E-9	7 / 53
t1.7	Polysaccharide metabolic process	1.79E-8	7 / 60
t1.8	Cellular carbohydrate metabolic process	3.76E-8	8 / 120
t1.9	Cellular polysaccharide catabolic process	1.09E-7	5 / 18
t1.10	Glycogen catabolic process	1.09E-7	5 / 17
t1.11	Glucose metabolic process	1.09E-7	8 / 144
t1.12	P catabolic process	1.09E-7	5 / 18
t1.13	Glucan catabolic process	1.09E-7	5 / 18
t1.14	Hexose metabolic process	2.12E-7	8 / 162
t1.15	Monosaccharide metabolic process	5.63E-7	8 / 185
t1.16	Energy reserve metabolic process	2.86E-6	7 / 141
t1.17	Cellular carbohydrate catabolic process	6.84E-6	5 / 41
t1.18	Energy derivation by oxidation of organic compounds	2.76E-4	7 / 279
t1.19	Single-organism carbohydrate catabolic process	4.17E-4	5 / 94
t1.20	Carbohydrate catabolic process	4.61E-4	5 / 97

Prediction of gene-associated functions was done by using the bioinformatic resource GenMANIA [28]. The input genes are illustrated in Fig. 1

patatin is an inducible storage protein. The protein encoded by the *PNPLA3* gene is an intracellular multifunctional enzyme that has both triacylglycerol lipase and acylglycerol O-acyltransferase activities [30] and shares domain and protein function with other members of the PNPLA family [31]. The protein has the serine lipase consensus motif GXSXG/A [32], which might have a role in the modulation of PNPLA3 by posttranslational modifications, including protein phosphorylation. In fact, PNPLA3 has many sites, including 21 Ser, 5 Thr, and 2 Tyr, with a high potential to be phosphorylated [33], being 134S close to the polymorphic I148M site. However, in spite of previous efforts, the functional meaning of the I148M variation remains to be established [31, 34–38].

GWAS on NAFLD and Translation of the Genomic Information into Clinical Practice

Whether or not the significant association between rs738409 and NAFLD and disease progression might be translated into clinical practice in terms of personalized medicine remains an open question. In such a scenario, it should be possible to make an individual risk assessment, and then the physician might be able to tailor a medical intervention (liver biopsy or disease therapy) based on the *PNPLA3* profile of the patient.

Unfortunately, despite the enormous enthusiasm for rs738409, current evidence shows that the role of the variant in predicting disease risk is not significantly improved compared with existing biomarkers of disease severity. For example, Kotronen and coworkers evaluated the performance of this SNP in predicting NAFLD by combining routine clinical and laboratory data and the rs738409 genotypes and observed a sensitivity of 86 % and a specificity of 71 % in the estimation of increased liver fat content [39]. Surprisingly, the addition of genetic information to the score improved the accuracy of the prediction by <1 %. Likewise, Francque et al. explored a set of routine and non-routine parameters, including ultrasound and genetic testing to predict the development of NASH in overweight or obese patients [40]. The authors observed that increased levels of ALT, fasting levels of C-peptide, and ultrasound steatosis scores had area under the receiver operating curve (AUROC) values of 0.854 and 0.823 for NASH development in the design and validation cohort, respectively [40]. In addition, authors observed that although the levels of cytokeratine18 and rs738409 correlated with the development of NASH, these did not add value to disease risk prediction [40]. Similarly, a recent study in a cohort of patients with medically complicated obesity showed that the probability of developing NASH was best predicted by a combination of four risk factors (the rs738409 G allele, CK-18 >145 IU/l, glucose >100 mg/dl, and C-reactive protein (CRP) >0.8 mg/dl): 82 % probability in the presence of all four risk factors versus 9 % in their absence) [41].

In conclusion, the incorporation of genetic tests in clinical practice is not that much more promising than the consideration of traditional and non canonical risk factors, which, when combined properly, have good predictive power [42, 43]. By contrast, following their observation that carriers of the GG genotype showed a twice higher independent risk for mortality, Hassan et al. reported that rs738409 may help predict poor survival and outcome of hepatocellular carcinoma [44]. Nevertheless, the importance of rs738409 in risk prediction remains unclear because, in the same cohort [44], other significant risk factors, including viral infection (HCV and HBV) and diabetes mellitus, also had significant predictive value. Indeed, because of the small effect associated with common variants, similarly to other complex diseases, genetic markers are still poor predictors [38].

Hence, the role of rs738409 in clinical decision making remains uncertain because there are no data to support that interventions should be restricted to carriers of the risk allele. In addition, clinicians who consider having their patients genotyped for the *PNPLA3* variant should carefully consider what type of information or recommendation could be returned to their patients, or the parents in the case of a pediatric population, because we do not yet have evidence as to whether patients carrying the risk variant will have a poor prognosis or even poor treatment response to any therapy. The bottom line is that to have an impact on predictive power, any variant should confer an odds ratio or risk of having the

disease of >10-200. Even in the best scenario and after combining many variants, such odds ratio may be reached. However, this would benefit a very small proportion of patients, at least in the absence of epistasis (a genetic interaction among the variants), a phenomenon that remains largely unexplored.

Epigenetic Changes, the Pathogenesis of NAFLD and MetSyn, and Potential Reversion of the Phenotype by Therapeutic Intervention

Epigenetic modifications have emerged as important mechanisms that modulate the pathogenesis of common diseases, including NAFLD, and present a potential explanation for the missing heredity. The main reason for this observation is that epigenetic changes are able to operate across the entire genome by regulating gene transcription and chromosome organization without affecting, by definition, the DNA sequence itself. More interestingly, epigenetic mechanisms are both highly regulated by environmental stimuli, including nutritional status, and highly dynamic.

One of the most common epigenetic modifications is DNA methylation, which occurs preferentially but not exclusively in the cytosine of the dinucleotide CpG. In normal conditions, the notable exceptions are the CpG-rich islands (regions typically 300–3000 bp in length with a high percentage content of CpG and C/G) present in the 5'- untranslated regions (5' UTR or promoters) of some genes. Nevertheless, epigenetic changes are not restricted to DNA methylation but also involve histone posttranslational modifications [45]. Therefore, although a comprehensive discussion of this issue is outside the scope of this review, it should be mentioned that many actors play a role in the epigenetic landscape; for example, DNA methyl transferases (DNMTs) and demethylases (TET and jumonji-domain protein families), and histone acetylases (HAT), deacetylases (HDAC), methylases, and demethylases [46].

Although epigenetics has attracted the genomic world in the last couple of years, the research agenda around NAFLD and epigenetics is very short, and our knowledge about epigenetic changes in human NAFLD is restricted to four studies. Our group showed the effect of epigenetic changes occurring in the fatty liver on the modulation of IR [8]. Further, we observed that the methylation status of the peroxisome proliferator-activated receptor gamma coactivator 1α (*PPARGC1A*) promoter is significantly associated with plasma fasting insulin levels and the homeostasis model assessment of IR (HOMA-IR) [8]. In addition, the methylation status of the *PPARGC1A* promoter was inversely correlated with the liver expression of the mRNA, suggesting that methylation of the explored CpG sites in the gene promoter efficiently repressed its transcriptional activity [8]. We also observed a complex interaction between the transcriptional activity of *PPARGC1A* and liver mitochondrial DNA copy number, which also had a direct impact on IR [8]. In our population, we showed that mitochondrial biogenesis was reduced in the liver of NAFLD patients and was associated with the peripheral IR and *PPARGC1A* promoter methylation status. A similar finding was observed in an experimental model of NAFLD [47]. It is worth noting that *PPARGC1A* is a master regulator of mitochondrial biogenesis and cell metabolism [48, 49]. Interestingly, many of these results were also observed in leukocyte DNA from adolescents [50] and umbilical cord DNA from small and large for gestational age in comparison with normal for gestational age newborns [51, 52]. Both extremes of fetal growth have been associated with MetSyn later in life, probably through epigenetic reprogramming of developmental and metabolism pathways [53].

An interesting study explored the pre- and post-bariatric changes in the methylation profile of nine genes coding for enzymes that regulate intermediate metabolism and insulin signaling in the liver of morbidly obese patients with NAFLD [54•]. The most remarkable finding of this study is that NAFLD-associated methylation changes were partially reversible by therapeutic intervention; for instance, the gene encoding protein-tyrosine phosphatase epsilon (*PTPRE*) showed both differential expression and differential methylation before and after bariatric surgery [54•]. Moreover, the authors observed that the insulin-like growth factor binding protein 2 (*IGFBP2*) locus was hypermethylated and its mRNA downregulated in NASH [54•].

Murphy and colleagues, who recently did global methylation profiling of liver samples of NAFLD patients at different stages of disease severity by using the Illumina HumanMethylation450 BeadChip platform, observed that patients with advanced NAFLD had a signature of differentially methylated CpG sites that allow discrimination between advanced versus mild disease [55•]. Indeed, the authors showed that advanced NAFLD has a relative hypomethylation state (11 % of 52,830 CpG sites) compared with mild NAFLD, specifically in genes associated with tissue repair; for instance, *FGFR2* (a fibroblast growth factor receptor family member), genes of the collagen (*COL1A1*, *COL1A2*, *COL4A1*, and *COL4A2*) and laminin families, and many chemokines [55•]. Of note, genes involved in pathways that generate methyl groups, including methylenetetrahydrofolate dehydrogenase 2 (*MTHFD2*) were significantly hypomethylated in advanced NAFLD [55•].

Finally, we recently described a novel disease mechanism associated with NAFLD progression that involves epigenetic changes of mitochondrial DNA (mtDNA) [7••]. In our study, we explored for the first time the status of cytosine methylation of liver mtDNA in target regions of the mtDNA genome. We observed that the methylation levels of NADH dehydrogenase 6 (MT-ND6), the gene that encodes for a key enzyme of complex 1 of the oxidative phosphorylation chain, were

higher in the liver of NASH patients and that there was a clear decrease in the protein level and changes in mitochondrial morphology, suggesting that the methylation status of this mitochondrial gene plays a role in the phenotypic switching from SS to NASH [7••]. To contrast with the hypothesis that epigenetic modifications might be reversible by intervention, we also explored whether the observed changes were associated with interventional programs. We observed that physical activity modulates the methylation status of MT-ND6 [7••].

In summary, epigenetics emerged as an interesting target of therapeutic intervention in chronic human diseases because it offers a unique framework of reversible mechanisms that modulate the cellular transcriptional machinery.

MicroRNAs and Modulation of the Transcriptional Machinery: Potential Epigenetic Modifiers in NAFLD by Fine-Tuning Modulation of Gene Transcription

MicroRNAs (miRNAs) are short noncoding RNAs that regulate gene expression at the posttranscriptional level. MiRNAs have emerged as powerful molecules in the transmission of information between cells. Moreover, genetic variation at the 3'-UTR gene containing a binding site for miRNAs has been associated with the regulation of gene transcription in human studies. For example, we observed that rs41318021 in the 3′-UTR of the human L-arginine transporter *SLC7A1* was significantly associated with arterial blood pressure in patients with NAFLD, suggesting a promising role for miRNAs in the epigenetic regulation of disease-associated traits in NAFLD patients [56].

Table 2 summarizes the results of human studies that have explored the expression of miRNAs either in circulation or in liver tissue. As expected, miRNA-122, the most abundant miRNA in the liver, is the most largely replicated miRNA deregulated in NAFLD; however, the mechanisms by which this miRNA operates in the modulation of the disease severity remain unclear. Evidence from in vitro silencing of miRNA-122 shows a time-regulated increase/decrease in the mRNA levels of lipogeneic genes, suggesting that miR-122 may operate by posttranslational regulation of mRNA maturation [61].

The exploration of enriched disease-associated pathways among miRNAs significantly deregulated in NAFLD by the bioinformatic resource TAM (a tool for annotation of human miRNAs; http://202.38.126.151/hmdd/tools/tam.html/) shows that miRNAs 122, 19a.b, 34a, and 21 are involved in the regulation of angiogenesis (p value<0.00004, Bonferoni p <0.014).

Finally, we used the resource DIANA-miRPath v2.1 (http://diana.imis.athena-innovation.gr/DianaTools/) to identify common disease pathways associated with the miRNAs mentioned in Table 2. Interestingly, we found a significantly predicted pathway associated with cancer (empirical p value=3.0 E^{-7}, false discovery correction) involving four miRNAs (hsa-miR-122-5p, hsa-miR-192-5p, hsa-miR-375, and hsa-miR-146b-5p) and targeting 46 genes, which might explain the role of the discovered miRNAs

t2.1 **Table 2** Role of miRNAs in human NAFLD: results from clinical studies about circulating and tissue expression

t2.2	Reference	Study design and sample size	miRNA: main findings
t2.3	Circulating miRNA		
t2.4	Cermelli et al. 2011 [57]	Observational study on patients with NAFLD proven by liver biopsy, no controls N=34	miR-34a and miR-122 represent noninvasive biomarkers for diagnosis and histologic disease severity
t2.5	Yamada et al. 2013 [58]	Population based, fatty liver explored by ultrasound scan N=430	miR-21, miR-34a, miR-122, and miR-451 were higher in participants with NAFLD miR-122 was correlated with severity of liver steatosis
t2.6	Min et al. 2012 [59]	Case-control study N=66	miR-34a increased in NAFLD
t2.7	Pirola et al. 2013 [60]	Case-control, 3 study phases (validation, replication, and tissue correlation), patients with NAFLD proven by liver biopsy N=209	miR-122, miR-192, miR-19a/b, and miR-375 increased in NAFLD and predict histologic disease severity
t2.8	Liver expression of miRNA		
t2.9	Cheung et al. 2008 [61]	Case-control study N=50	miR-122 level was significantly decreased in subjects with NASH. miR-34a and miR-146b levels were significantly increased in subjects with NASH
t2.10	Pirola et al. 2013 [60]	Case-control, patients with NAFLD proven by liver biopsy N=65	miR-122 level was 10-fold decreased in subjects with NASH.

associated with NAFLD in the progression of the disease and hepatocarcinogenesis. Interestingly, many of the predicted targets of these miRNAs have been shown to be dysregulated in NASH versus simple steatosis [62].

Conclusions

- rs738409 is the most consistently replicated SNP worldwide that influences the genetic risk of NAFLD and disease progression.
- GWAS on NAFLD around the world have replicated the *PNPLA3* signal and uncovered new gene variants, for which replication and functional analysis are needed to better understand their role in the pathogenesis of NAFLD.
- The incorporation of genetic testing into clinical practice for predicting NAFLD progression or determining disease intervention remains incipient; large and well-conducted clinical trials are needed to determine its real advantage and performance in comparison with classical (ALT, CK-18, etc.) noninvasive biomarkers, algorithms including simple clinical characteristics, or the gold standard (liver biopsy).
- Epigenetic changes are promising molecular mechanisms for explaining disease pathogenesis. The dynamic nature of epigenetic modifications is an attractive target for therapeutic intervention because of the potential reversibility of the liver changes observed in NAFLD after physical activity or bariatric surgery and even after the administration of existing drugs or natural compounds. Mitochondrial epigenetics has emerged as an interesting mechanism for explaining disease transition from simple steatosis to NASH.
- Noncoding miRNAs are deregulated in the circulation and in the liver of NAFLD patients and might explain the predisposition to liver cancer.

Acknowledgments This study was partially supported by grants PICT 2010-0441 and PICT 2012-0159 (Agencia Nacional de Promoción Científica y Tecnológica), and UBACYT CM04 (Universidad de Buenos Aires). SS and CJP belong to Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Compliance with Ethics Guidelines

Conflict of Interest Silvia Sookoian declares no conflicts of interest.

Carlos J. Pirola is a paid board member, receives honoraria, and receives travel/accomodation expenses from Merck Sharp and Dohm.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1. Sookoian S, Pirola CJ. The genetic epidemiology of nonalcoholic fatty liver disease: toward a personalized medicine. Clin Liver Dis. 2012;16(3):467–85.
2. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2008;40(12):1461–5.
3. Rodriguez B, Torres DM, Harrison SA. Physical activity: an essential component of lifestyle modification in NAFLD. Nat Rev Gastroenterol Hepatol. 2012;9(12):726–31.
4. Eckard C, Cole R, Lockwood J, Torres DM, Williams CD, Shaw JC, et al. Prospective histopathologic evaluation of lifestyle modification in nonalcoholic fatty liver disease: a randomized trial. Ther Adv Gastroenterol. 2013;6(4):249–59.
5. Cortez-Pinto H, Machado M. Impact of body weight, diet and lifestyle on nonalcoholic fatty liver disease. Expert Rev Gastroenterol Hepatol. 2008;2(2):217–31.
6. Neuschwander-Tetri BA. Lifestyle modification as the primary treatment of NASH. Clin Liver Dis. 2009;13(4):649–65.
7.•• Pirola CJ, Gianotti TF, Burgueno AL, Rey-Funes M, Loidl CF, Mallardi P, et al. Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. Gut. 2013;62(9):1356–63. *This is the first study showing epigenetic changes in mitochondrial DNA having an impact on a human disease.*
8. Sookoian S, Rosselli MS, Gemma C, Burgueno AL, Fernandez GT, Castano GO, 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. *This is the first human study showing that changes in liver DNA methylation of PGC1A modulate insulin resistance.*
9. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. Gastroenterology. 2012;142(7):1592–609.
10. Sookoian S, Castano GO, Pirola CJ. Modest alcohol consumption decreases the risk of non-alcoholic fatty liver disease: a meta-analysis of 43 175 individuals. Gut. 2013. doi:10.1136/gutjnl-2013-305718.
11. Brunt EM. Pathology of fatty liver disease. Mod Pathol. 2007;20 Suppl 1:S40–8.
12. Sookoian S, Pirola CJ. Systems biology elucidates common pathogenic mechanisms between nonalcoholic and alcoholic-fatty liver disease. PLoS One. 2013;8(3):e58895.
13. Stickel F, Hampe J. Genetic determinants of alcoholic liver disease. Gut. 2012;61(1):150–9.
14. Sookoian S, Castano GO, Burgueno AL, Gianotti TF, Rosselli MS, Pirola CJ. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res. 2009;50(10):2111–6.
15. 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.

16. Chalasani N, Guo X, Loomba R, Goodarzi MO, Haritunians T, Kwon S, et al. Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. Gastroenterology. 2010;139(5):1567–76, 1576.
17. Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet. 2011;7(3): e1001324.
18. Feitosa MF, Wojczynski MK, North KE, Zhang Q, Province MA, Carr JJ, et al. The ERLIN1-CHUK-CWF19L1 gene cluster influences liver fat deposition and hepatic inflammation in the NHLBI Family Heart Study. Atherosclerosis. 2013;228(1):175–80.
19. Kitamoto T, Kitamoto A, Yoneda M, Hyogo H, Ochi H, Nakamura T, 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. Hum Genet. 2013;132(7):783–92.
20. Kawaguchi T, Sumida Y, Umemura A, Matsuo K, Takahashi M, Takamura T, et al. Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. PLoS One. 2012;7(6):e38322.
21. Yuan X, Waterworth D, Perry JR, Lim N, Song K, Chambers JC, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet. 2008;83(4):520–8.
22. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, van der Harst P, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet. 2011;43(11):1131–8.
23. Palmer ND, Musani SK, Yerges-Armstrong LM, Feitosa MF, Bielak LF, Hernaez R, et al. Characterization of european ancestry nonalcoholic fatty liver disease-associated variants in individuals of african and hispanic descent. Hepatology. 2013;58(3):966–75.
24. Hernaez R, McLean J, Lazo M, Brancati FL, Hirschhorn JN, Borecki IB, et al. Association between variants in or near PNPLA3, GCKR, and PPP1R3B with ultrasound-defined steatosis based on data from the third national health and nutrition examination survey. Clin Gastroenterol Hepatol. 2013;11(9):1183–90.
25. Trepo E, Nahon P, Bontempi G, Valenti L, Falleti E, Nischalke HD, et al. Association between the PNPLA3 (rs738409 C>G) variant and hepatocellular carcinoma: evidence from a meta-analysis of individual participant data. Hepatology. 2013. doi:10.1002/hep.26767.
26. Gerhard GS, Chu X, Wood GC, Gerhard GM, Benotti P, Petrick AT, et al. Next-generation sequence analysis of genes associated with obesity and nonalcoholic fatty liver disease-related cirrhosis in extreme obesity. Hum Hered. 2013;75(2–4):144–51.
27. Cefalu AB, Pirruccello JP, Noto D, Gabriel S, Valenti V, Gupta N, et al. A novel APOB mutation identified by exome sequencing cosegregates with steatosis, liver cancer, and hypocholesterolemia. Arterioscler Thromb Vasc Biol. 2013;33(8):2021–5.
28. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, 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.
29. Mignery GA, Pikaard CS, Park WD. Molecular characterization of the patatin multigene family of potato. Gene. 1988;62(1):27–44.
30. Jenkins CM, Mancuso DJ, Yan W, Sims HF, Gibson B, Gross RW. 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.
31. Sookoian S, Pirola CJ. PNPLA3, the triacylglycerol synthesis/hydrolysis/storage dilemma, and nonalcoholic fatty liver disease. World J Gastroenterol. 2012;18(42):6018–26.
32. Kim TS, Sundaresh CS, Feinstein SI, Dodia C, Skach WR, Jain MK, et al. Identification of a human cDNA clone for lysosomal type Ca2+-independent phospholipase A2 and properties of the expressed protein. J Biol Chem. 1997;272(4):2542–50.
33. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol. 1999;294(5):1351–62.
34. He S, McPhaul C, Li JZ, Garuti R, Kinch L, Grishin NV, et al. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. J Biol Chem. 2010;285(9):6706–15.
35. Li JZ, Huang Y, Karaman R, Ivanova PT, Brown HA, Roddy T, et al. Chronic overexpression of PNPLA3I148M in mouse liver causes hepatic steatosis. J Clin Invest. 2012;122(11):4130–44.
36. Perttila J, Huaman-Samanez C, Caron S, Tanhuanpaa K, Staels B, Yki-Jarvinen H, et al. PNPLA3 is regulated by glucose in human hepatocytes, and its I148M mutant slows down triglyceride hydrolysis. Am J Physiol Endocrinol Metab. 2012;302(9):E1063–9.
37. Pirazzi C, Adiels M, Burza MA, Mancina RM, Levin M, Stahlman M, et al. Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. J Hepatol. 2012;57(6):1276–82.
38. Sookoian S, Pirola CJ. PNPLA3, the history of an orphan gene of the potate tuber protein family that found an organ: the Liver. Hepatology. 2013. doi:10.1002/hep.26895.
39. Kotronen A, Peltonen M, Hakkarainen A, Sevastianova K, Bergholm R, Johansson LM, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. Gastroenterology. 2009;137(3):865–72.
40. Francque SM, Verrijken A, Mertens I, Hubens G, Van ME, Pelckmans P, et al. Noninvasive assessment of nonalcoholic fatty liver disease in obese or overweight patients. Clin Gastroenterol Hepatol. 2012;10(10):1162–8.
41. Guichelaar MM, Gawrieh S, Olivier M, Viker K, Krishnan A, Sanderson S, et al. Interactions of allelic variance of PNPLA3 with nongenetic factors in predicting nonalcoholic steatohepatitis and nonhepatic complications of severe obesity. Obesity (Silver Spring). 2013;21(9):1935–41.
42. Adams LA, Feldstein AE. Non-invasive diagnosis of nonalcoholic fatty liver and nonalcoholic steatohepatitis. J Dig Dis. 2011;12(1): 10–6.
43. Sookoian S, Castano G, Burgueno AL, Gianotti TF, Rosselli MS, Pirola CJ. 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.
44. Hassan MM, Kaseb A, Etzel CJ, El-Serag H, Spitz MR, Chang P, et al. Genetic variation in the PNPLA3 gene and hepatocellular carcinoma in USA: Risk and prognosis prediction. Mol Carcinog. 2013;52 Suppl 1:139–47.
45. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature. 2004;429(6990):457–63.
46. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009;10(5):295–304.
47. Burgueno AL, Cabrerizo R, Gonzales MN, Sookoian S, Pirola CJ. Maternal high-fat intake during pregnancy programs metabolic-syndrome-related phenotypes through liver mitochondrial DNA copy number and transcriptional activity of liver PPARGC1A. J Nutr Biochem. 2013;24(1):6–13.
48. Ventura-Clapier R, Garnier A, Veksler V. Transcriptional control of mitochondrial biogenesis: the central role of PGC-1alpha. Cardiovasc Res. 2008;79(2):208–17.
49. Sookoian S, Pirola CJ. DNA methylation and hepatic insulin resistance and steatosis. Curr Opin Clin Nutr Metab Care. 2012;15(4): 350–6.

50. Gianotti TF, Sookoian S, Dieuzeide G, Garcia SI, Gemma C, Gonzalez CD, et al. A decreased mitochondrial DNA content is related to insulin resistance in adolescents. Obesity. 2008;16(7):1591–5.
51. Gemma C, Sookoian S, Alvarinas J, Garcia SI, Quintana L, Kanevsky D, et al. Mitochondrial DNA depletion in small- and large-for-gestational-age newborns. Obesity (Silver Spring). 2006;14(12):2193–9.
52. Gemma C, Sookoian S, Alvarinas J, Garcia SI, Quintana L, Kanevsky D, et al. Maternal pregestational BMI is associated with methylation of the PPARGC1A promoter in newborns. Obesity (Silver Spring). 2009;17(5):1032–9.
53. Sookoian S, Gianotti TF, Burgueno AL, Pirola CJ. Fetal metabolic programming and epigenetic modifications: a systems biology approach. Pediatr Res. 2013;73(4 Pt 2):531–42.
54.• Ahrens M, Ammerpohl O, von Schönfels W, Kolarova J, Bens S, Itzel T, et al. DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. Cell Metab. 2013;18(2):296–302. *Human study showing the potential reversibility of epigenetic changes after therapeutic intervention.*
55.• Murphy SK, Yang H, Moylan CA, Pang H, Dellinger A, Abdelmalek MF, et al. Relationship between methylome and transcriptome in patients with nonalcoholic Fatty liver disease. Gastroenterology. 2013;145(5):1076–87. *This study explored methylation marks and its relationship with the disease severity.*
56. Pirola CJ, Gianotti TF, Castano GO, Sookoian S. Circulating MicroRNA-122 signature in nonalcoholic fatty liver disease and cardiovascular disease: a new endocrine system in metabolic syndrome. Hepatology. 2013;57(6):2545–7.
57. Cermelli S, Ruggieri A, Marrero JA, Ioannou GN, Beretta L. Circulating microRNAs in patients with chronic hepatitis C and non-alcoholic fatty liver disease. PLoS One. 2011;6(8):e23937.
58. Yamada H, Suzuki K, Ichino N, Ando Y, Sawada A, Osakabe K, et al. Associations between circulating microRNAs (miR-21, miR-34a, miR-122 and miR-451) and non-alcoholic fatty liver. Clin Chim Acta. 2013;424:99–103.
59. Min HK, Kapoor A, Fuchs M, Mirshahi F, Zhou H, Maher J, et al. Increased hepatic synthesis and dysregulation of cholesterol metabolism is associated with the severity of nonalcoholic fatty liver disease. Cell Metab. 2012;15(5):665–74.
60. Pirola CJ, Fernandez GT, Burgueno AL, Castano G, Sookoian S. Circulating micro-RNA Profile in nonalcoholic fatty liver disease: potential biomarkers and its role in the modulation of the metabolic syndrome? Hepatology. 2013;58(suppl):247A.
61. Cheung O, Puri P, Eicken C, Contos MJ, Mirshahi F, Maher JW, et al. Nonalcoholic steatohepatitis is associated with altered hepatic MicroRNA expression. Hepatology. 2008;48(6):1810–20.
62. Sookoian S, Gianotti TF, Rosselli MS, Burgueno AL, Castano GO, Pirola CJ. Liver transcriptional profile of atherosclerosis-related genes in human nonalcoholic fatty liver disease. Atherosclerosis. 2011;218(2):378–85.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES.

Q1. Please confirm that this information is correct. Also note that, in accordance with PubMed rules, information on studies with human or animal subjects must be included not only for original information included in this submitted paper, but also for papers cited in references for which one of the authors was also an author.

Q2. In reference 54, Author “von SW” has been changed to von Schönfels W. Please check if correct.