

RAPID COMMUNICATION

## Common genetic variations in *CLOCK* transcription factor are associated with nonalcoholic fatty liver disease

Silvia Sookoian, Gustavo Castaño, Carolina Gemma, Tomas Fernández Gianotti, Carlos Jose Pirola

Silvia Sookoian, Carolina Gemma, Tomas Fernández Gianotti, Carlos Jose Pirola, Instituto de Investigaciones Medicas, A. Lanari. Universidad de Buenos Aires, CONICET, Combatiente de Malvinas 3150, 1427- Ciudad Autonoma de Buenos Aires, Argentina

Silvia Sookoian, Gustavo Castaño, Consejo de Investigación, GCBA, Buenos Aires, Argentina

Supported by Grant B119 (Universidad de Buenos Aires), PICT 25920 (Agencia Nacional de Promoción Científica y Tecnológica) and PIP 5195 (Consejo Nacional de Investigaciones Científicas y Técnicas). SS, CG and CJP belong to Consejo Nacional de Investigaciones Científicas y Técnicas

Correspondence to: Silvia Sookoian, MD, PhD, Instituto de Investigaciones Medicas, A. Lanari. Universidad de Buenos Aires, CONICET, Combatiente de Malvinas 3150, 1427- Ciudad Autonoma de Buenos Aires, Argentina. ssookoian@intramed.net  
Telephone: +54-11-45148701-167 Fax: +54-11-45238947

Received: 2007-03-26

Accepted: 2007-04-27

susceptibility to NAFLD and disease severity.

© 2007 WJG. All rights reserved.

**Key words:** Circadian rhythm; Circadian locomoter output cycles protein kaput; Gene variants; Haplotypes; Fatty liver; Liver disease; Obesity; Nonalcoholic steatohepatitis

Sookoian S, Castaño G, Gemma C, Gianotti TF, Pirola CJ. Common genetic variations in *CLOCK* transcription factor are associated with nonalcoholic fatty liver disease. *World J Gastroenterol* 2007; 13(31): 4242-4248

<http://www.wjgnet.com/1007-9327/13/4242.asp>

### Abstract

**AIM:** To investigate the role of gene variants and derived haplotypes of the *CLOCK* transcription factor in nonalcoholic fatty liver disease (NAFLD) and their relation with the disease severity.

**METHODS:** A total of 136 patients with NAFLD and 64 healthy individuals were studied. Liver biopsy was performed in 91 patients. Six tag SNPs showing a minor allele frequency > 10% (rs1554483 C/G; rs11932595 A/G; rs4580704 C/G; rs6843722 A/C; rs6850524 C/G and rs4864548 A/G) encompassing 117 kb of chromosome 4 and representing 115 polymorphic sites ( $r^2 > 0.8$ ) were genotyped.

**RESULTS:** rs11932595 and rs6843722 showed significant associations with NAFLD (empiric  $P = 0.0449$  and  $0.023$ , respectively). A significant association was also observed between clinical or histologic spectrum of NAFLD and rs1554483 (empiric  $P = 0.0399$ ), rs6843722 (empiric  $P = 0.0229$ ) and rs6850524 (empiric  $P = 0.00899$ ) and between fibrosis score and rs1554483 (empiric  $P = 0.02697$ ), rs6843722 (empiric  $P = 0.01898$ ) and rs4864548 (empiric  $P = 0.02697$ ). Test of haplotypic association showed that *CLOCK* gene variant haplotypes frequencies in NAFLD individuals significantly differed from those in controls (empiric  $P = 0.0097$ ).

**CONCLUSION:** Our study suggests a potential role of the *CLOCK* polymorphisms and their haplotypes in

### INTRODUCTION

During the past 20 years, obesity has reached worldwide epidemic proportions, not only among adults but also in young people, thus becoming a major contributor to the global health burden imposed by chronic disease and disability<sup>[1,2]</sup>. Nonalcoholic fatty liver disease (NAFLD) is one of the most common abnormalities observed in obese persons and refers to a wide spectrum of liver diseases<sup>[3,4]</sup>, ranging from fatty liver alone to nonalcoholic steatohepatitis (NASH) with evidence of liver cell injury, a mixed inflammatory lobular infiltrate, and variable fibrosis<sup>[5]</sup>. Estimates based on imaging and autopsy studies suggest that about 20% to 30% of the adult population in Western countries have excess fat accumulation in the liver<sup>[6]</sup>.

The pathogenesis of NAFLD is multifactorial<sup>[6]</sup> and, as a complex disease, the disorder develops from the interplay between genes and the environment<sup>[7,8]</sup>. While it is well known that environmental risk factors influence the development and progression of NAFLD<sup>[9]</sup>, the contribution of genetic variation to the disease predisposition is still inconclusive despite the fact that several genes have been suggested as potential candidates for advanced NAFLD progression<sup>[10-14]</sup>.

In addition, the ectopic accumulation of fat in the liver has been strongly associated with metabolic factors. For instance, insulin resistance, intra-abdominal fat and body fat distribution appear to play an important role in the pathogenesis of NAFLD<sup>[15]</sup>. In fact, the metabolic syndrome is a strong predictor of NAFLD<sup>[16]</sup>.

In mammals, many physiological processes, including energy metabolism and food intake, are driven by a

circadian timing system comprising a master pacemaker in the suprachiasmatic nuclei of the hypothalamus and peripheral oscillators in most body cells<sup>[17,18]</sup>. Besides, many transcripts that participate in common metabolic pathways, such as metabolism of glucose, cholesterol and fatty acids, show circadian rhythmicity<sup>[19,20]</sup>.

A recent report by Turek *et al.*<sup>[21]</sup> showed that homozygous circadian *CLOCK* (circadian locomotor output cycles protein kaput) mutant mice had a greatly altered diurnal feeding rhythm, were hyperphagic and obese, and developed a metabolic syndrome with hyperleptinemia, hyperlipidemia, hypoinsulinemia, hyperglycemia, and hepatic steatosis.

Given the results of the aforementioned studies, which showed that altered circadian rhythmicity resulted in pathophysiological changes resembling metabolic syndrome and fatty liver disease, the aim of this study was to investigate the role of gene variants and their predicted haplotypes of the linkage disequilibrium (LD) block of the *CLOCK* gene in NAFLD in a candidate gene case-control association study. Additionally, we tested the hypothesis of a relation between the gene variants and disease severity.

## MATERIALS AND METHODS

### Subjects

A total of 136 consecutive unrelated patients of self-reported European ancestry with features of NAFLD (36 males and 100 females), including ultrasonographic (USG) examinations suggestive of fatty infiltration<sup>[22]</sup> performed in all the cases by the same operator, were included in this study. Secondary causes of steatosis, including alcohol abuse ( $\geq 30$  g alcohol daily for men and  $\geq 20$  g for women), total parenteral nutrition, hepatitis B and hepatitis C virus infection, and the use of drugs known to precipitate steatosis, were excluded in all cases. By using standard clinical and laboratory evaluation as well as liver biopsy features when applicable, autoimmune liver disease, metabolic liver disease, Wilson's disease, and  $\alpha$ -1-antitrypsin deficiency were likewise ruled out in all patients.

For the evaluation of the disease severity, NAFLD cases were classified as follows: fatty liver with persistently normal liver function test (FL-NLFT); fatty liver with persistently abnormal liver function test (FL-ALFT); and NASH proven through biopsy as described below. Patients were defined to have abnormal liver function test in the presence of at least one of the following criteria: (1) elevated serum alanine (ALT) and/or aspartate aminotransferase (AST), defined as  $> 41$  U/L; (2) gamma-glutamyl-transferase (GGT)  $> 50$  U/L; and (3) alkaline phosphatase (FA)  $> 250$  UI/L.

Additionally, 64 healthy individuals (22 males and 42 females) with the same demographic background and who underwent the annual health examination during the same study period were included in the study as a control group. All healthy controls were subjected to USG examination. None of them evidenced fatty change, biochemical abnormalities or features indicative of metabolic syndrome.

### Physical, anthropometric and biochemical evaluation

Health examinations included anthropometric measurements, a questionnaire on health-related behaviours, and

biochemical determinations. Body mass index (BMI) was calculated as weight/height<sup>2</sup> (kg/m<sup>2</sup>) and was used as the index for relative weight. Additionally, waist and hip circumferences were also assessed. Blood was drawn from fasting subjects who had lain in a supine resting position for at least 30 min. Serum insulin, total cholesterol, HDL and LDL-cholesterol, triglycerides, plasma glucose and liver function tests were measured by standard clinical laboratory techniques. Homeostasis Model Assessment (HOMA) was used to evaluate an insulin resistance index and was calculated as fasting serum insulin ( $\mu$ U/mL)  $\times$  fasting plasma glucose (mmol/L)/22.5. Insulin resistance was defined as HOMA index  $\geq 2.4$ <sup>[23]</sup>. Elevated blood pressure was defined as systolic arterial blood pressure (SABP)  $\geq 130$  mmHg and/or diastolic arterial blood pressure (DABP)  $\geq 85$  mmHg or receipt of anti-hypertensive medications. All the investigations performed in this study were conducted in accordance with the guidelines of the Declaration of Helsinki. Written consent from individuals was obtained in accordance with the procedures approved by the Ethical Committee of our institution.

### Liver biopsies and histopathological evaluation

A percutaneous liver biopsy (LB) was performed in 91 patients with fatty changes in the liver on USG plus persistently abnormal liver function tests (in at least 3 different determinations in a follow-up for 6 months period). LB was performed under ultrasound guidance with a modified 1.4-mm diameter Menghini needle (Hepafix, Braun, Germany) on an outpatient basis. Liver biopsy specimens were routinely fixed in 40 g/L formaldehyde (pH 7.4), embedded in paraffin and stained with hematoxylin and eosin, Masson trichrome and silver impregnation for reticular fibers. All biopsies were read by the same liver pathologist who was blinded to patient details. The diagnosis of NASH was confirmed by liver histology, and grading of necroinflammatory activity or staging of fibrosis was scored according to the system developed by Brunt *et al.*<sup>[24]</sup>. NASH was defined as steatosis plus any stage of fibrosis or as steatosis plus lobular inflammation plus ballooning degeneration<sup>[6]</sup>.

### Genotype and haplotype analysis

The genetic analyses were done on genomic DNA extracted from white blood cells by a standard method as previously described<sup>[25]</sup>.

In order to assess the contribution of Clock gene variants to NAFLD, we selected tag SNPs (tSNPs) by using an aggressive tagging approach to construct single marker and multi-marker tests to capture alleles of interest<sup>[26]</sup> and the phase II genotyping data from the HapMap project for Caucasians from the CEU dataset with a minor allele frequency (MAF)  $\geq 0.10$  and a minimum  $r^2$  of 0.8.

Genotyping was performed by a high-throughput genotyping method involving PCR amplification of genomic DNA with two-tailed allele-specific primers that introduce priming sites for universal energy-transfer-labeled primers as previously described<sup>[27]</sup>.

### Statistical analysis

The PLINK software was used for assessing association

**Table 1 Clinical and biochemical characteristics of the study individuals**

Variables	Control group	NAFLD patients	Empiric P value
Number of subjects	64	136	
Age, yr	46.14 ± 1.64	55.58 ± 1.06	0.001
Smoking habit, cigarettes/d	1.58 ± 0.52	1.30 ± 0.75	NS
Physical activity, h/wk	1.4 ± 0.30	1.4 ± 0.32	NS
Drinking habits, g alcohol/d	1.72 ± 0.64	0.75 ± 0.42	NS
BMI, kg/m <sup>2</sup>	25.61 ± 0.64	36.02 ± 3.11	0.001
Waist circumference, cm	84.13 ± 2.14	103.07 ± 1.42	0.001
Waist-hip ratio	0.84 ± 0.01	0.91 ± 0.007	0.001
SABP, mmHg	121.64 ± 1.87	124.05 ± 2.11	NS
DABP, mmHg	75.86 ± 1.32	78.38 ± 1.49	NS
Fasting plasma glucose, mmol/L	4.73 ± 0.07	5.74 ± 0.22	0.050
Fasting plasma insulin, pmol/L	46.1 ± 3.5	94.5 ± 6.9	0.001
HOMA Index	1.42 ± 0.12	3.44 ± 0.30	0.001
Total cholesterol, mmol/L	5.59 ± 0.17	5.44 ± 0.15	NS
HDL cholesterol, mmol/L	1.27 ± 0.09	1.20 ± 0.05	NS
LDL-cholesterol, mmol/L	3.22 ± 0.29	3.11 ± 0.14	NS
Uric acid, μmol/L	220 ± 51	244 ± 27	NS
Triglycerides, mmol/L	1.74 ± 0.13	2.02 ± 0.13	NS
ALT, UI/L	24.37 ± 5.2	49.25 ± 5.32	0.080
AST, UI/L	26.4 ± 2.0	36.72 ± 2.27	0.080
γGT, UI/L	36.50 ± 6.8	53.69 ± 5.32	NS
AF, UI/L	221.87 ± 13.0	228.90 ± 11.2	NS

NAFLD: Non-alcoholic fatty liver disease; BMI: Body mass index; SABP and DABP: Systolic and diastolic arterial blood pressure, respectively; HOMA: Homeostatic model assessment; ALT and AST: Serum alanine and aspartate aminotransferase, respectively; γGT: Gamma-glutamyl-transferase; AF: Alkaline phosphatase. Results are expressed as mean ± SE. Nominal P value stands for statistical significance using Student's t test after log variable transformation. Empiric P value stands for statistical significance adjusted for multiple testing by Bonferoni correction. NS: Non significant.

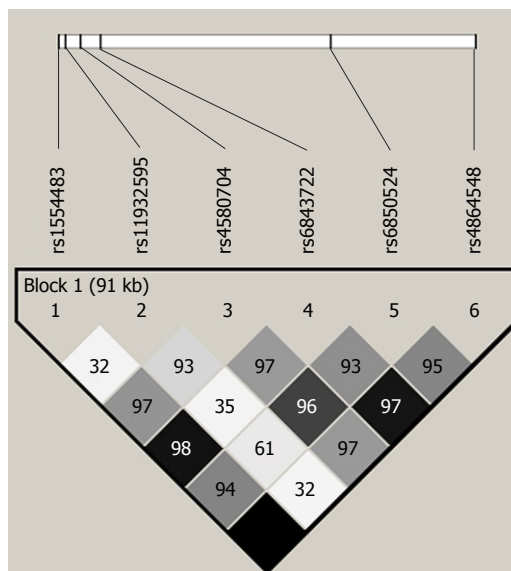
between SNPs and affection status and quantitative traits as well as for testing Hardy-Weinberg equilibrium and LD measures. SNP haplotype analysis was performed by both WHAP<sup>[28]</sup> and Haploview<sup>[29]</sup> softwares. These tools were also used to obtain haplotype frequencies. Control for multiple testing was done by permutation testing of individual traits to obtain an empirical P-value.

Differences in genotype frequencies between cases and controls were analyzed using PLINK and WHAP softwares as previously described. Multi-marker haplotype test was performed by Haploview software<sup>[29]</sup>.

Phenotypic quantitative data were expressed as mean ± SE. For univariate analysis, differences between groups were assessed by Student's t test on log-transformed variables. The conservative Bonferroni procedure was used for multiple testing corrections. We used the CSS/Statistica program package, StatSoft V 6.0 (Tulsa, USA) to perform these analyses.

## RESULTS

Table 1 shows the clinical features, anthropometric variables and laboratory findings of the cases and controls. NAFLD patients were older and had most of the risk factors of the metabolic syndrome: elevated BMI, waist-hip ratio, waist circumference, fasting insulin level, and HOMA index. Of 136 patients, 45 (33%) were classified as having FL-NLFT, 35 (25.7%) as FL-ANFT, and 56 (41.2%)



**Figure 1** Linkage disequilibrium plot across the *CLOCK* gene. The horizontal white line depicts the 117-kb DNA segment of chromosome 4q12 analyzed in our sample. The 6 tagSNPs locations are indicated by hatch marks. A linkage disequilibrium plot is depicted in the bottom part of the figure. Each diamond represents the magnitude of LD for a single pair of markers, with colors indicating strong LD (black,  $r^2 = 1.0$ ) and no LD (white,  $r^2 = 0$ ) as the extremes (different gray tones indicate intermediate LD). Numbers inside the diamonds stand for D' values.

**Table 2** Tag single nucleotide polymorphisms of the *CLOCK* gene genotyped in the study

NCBI SNP reference <sup>1</sup>	Location in the <i>CLOCK</i> gene	Heterozygosity	dsNP allele	Minor allele	MAF
rs1554483	Intron 12	0.494	C/G	G	0.491
rs11932595	Intron 10	0.478	A/G	G	0.288
rs4580704	Intron 8	0.395	C/G	G	0.311
rs6843722	Intron 7	0.483	A/C	C	0.476
rs6850524	Intron 1	0.490	C/G	C	0.3557
rs4864548	Upstream	0.490	G/A	A	0.478

MAF: Minor allele frequency (within the study); *CLOCK*: Circadian locomotor output cycles protein kaput. <sup>1</sup>Single nucleotide polymorphisms on NCBI Reference Assembly.

as NASH proven through biopsy.

### *CLOCK* gene variants

To diminish the burden of genotyping the multiple variants present in the *CLOCK* gene, we selected 6 tagSNPs showing a minor allele frequency (MAF) > 10% (rs1554483 C/G; rs11932595 A/G; rs4580704 C/G; rs6843722 A/C; rs6850524 C/G and rs4864548 G/A) encompassing 115 polymorphic sites with an  $r^2 > 0.8$  considering the HapMap project data (Figure 1); Table 2 illustrates the tSNPs description. Test results from the Tagger algorithm showed the aforementioned single-marker tags and 8 multi-marker tags capturing all SNPs of MAF ≥ 5% (Table 3). The distribution of the genotypes was in Hardy-Weinberg equilibrium (data not shown).

In univariate analysis, after the multiple comparison correction by permutation tests, there were significant

**Table 3** Association of single-tagSNP and multi-marker tag SNP tests (haplotypes) in *CLOCK* with NAFLD

Single tagSNP	Associated allele	Frequency	OR (95% CI)	Empiric P value
rs1554483	G	0.491	1.51 (0.93-2.45)	0.0599
rs11932595	A	0.712	1.65 (1.10-2.47)	0.0449
rs4580704	C	0.688	1.03 (0.68-1.55)	0.99
rs6843722	C	0.476	1.70 (1.05-2.76)	0.0239
rs6850524	G	0.643	1.40 (0.85-2.29)	0.1458
rs4864548	A	0.478	1.54 (0.95-2.50)	0.0779
Multi-marker tagSNP test	Associated haplotype	Frequency	OR (95% CI)	Empiric P value
rs11932595, rs6850524	AC	0.315	0.88 (1.55-1.41)	0.91
	GC	0.039	0.32 (0.10-0.94)	0.099
	CA	0.477	1.55 (0.99-2.44)	0.1066
rs4580704, rs4864548	GG	0.307	0.90 (0.56-1.46)	0.8764
	CG	0.212	0.62 (0.37-1.07)	0.119
	CA	0.213	0.58 (0.34-0.98)	0.0792
rs4580704, rs6843722	CA	0.213	0.58 (0.34-0.98)	0.0792
rs4580704, rs6850524	GC	0.300	0.90 (0.56-1.47)	0.939
rs6850524, rs4864548	GG	0.171	0.81 (0.46-1.45)	0.806

Empiric *P* values were obtained by adjusting nominal *P* values for multiple testing by permutation test.

differences in the allele frequency of the rs11932595 and rs6843722 between the control group and NAFLD patients (empiric *P* = 0.0449 and 0.0239, respectively). The test results for the multi-marker analysis showed that rs11932595/rs6850524 (haplotype GC) and rs4580704/rs6843722 (haplotype CA) were significantly associated with NAFLD (nominal *P* = 0.0424 and 0.0373, respectively). However, permutation testing to correct for the multiple hypothesis yielded a non-significant empirical *P* value for these comparisons (*P* = 0.099 and 0.0792, respectively) (Table 3).

When we tested the hypothesis of a relation between the gene variants and the clinical and histologic spectrum of NAFLD (disease severity), a significant association was observed with rs1554483 (empiric *P* = 0.039), rs6843722 (empiric *P* = 0.0229) and rs6850524 (empiric *P* = 0.00899). Moreover, in the multi-marker analysis, a significant association was observed for rs11932595/rs6850524 (haplotype AC, empiric *P* = 0.04858 and haplotype GC, empiric *P* = 0.0143) and for rs4580704/rs4864548 (haplotype CA, empiric *P* = 0.0248).

Besides, when we analyzed the allele frequencies of tSNPs in patients with NASH, we observed a significant association with the overall fibrosis score of the disease and rs1554483 (empiric *P* = 0.02697), rs6843722 (empiric *P* = 0.01898) and rs4864548 (empiric *P* = 0.02697). However, no association was observed between the necroinflammatory grade and the tSNPs.

Using the WHAP software, we further evaluated the LD pattern for the aforementioned tSNPs. The estimation of haplotypes from genotype data showed 5 possible combinations covering 94.8% of the common haplotypes with a frequency higher than 0.01 (Table 4). In the *omnibus* test of haplotypic association, we observed that *CLOCK* gene variant haplotypes frequencies in NAFLD individuals significantly differed from those in control subjects (empiric *P* = 0.0097). Haplotypes GACCGA and CGCACG explain

**Table 4** Estimated haplotype frequency distributions of *CLOCK* gene SNPs in the control group and in NAFLD patients

Haplotype	Control group (n = 128)	NAFLD patients (n = 272)	OR (95% CI)	Empiric P level
GACCGA	0.275	0.433	1.82 (1.22-2.70)	0.00562
CAGACG	0.363	0.309	0.89 (0.54-1.49)	0.316
CGCAGG	0.160	0.147	0.84 (0.46-1.53)	0.737
GGCCGA	0.118	0.088	0.73 (0.35-1.54)	0.371
CGCACG	0.084	0.024	0.27 (0.09-0.83)	0.00902

NAFLD: Non-alcoholic fatty liver disease. Haplotypes are composed by variants of rs1554483, rs11932595, rs4580704, rs6843722, rs6850524 and 4864548 in that order. *n*: Number of chromosomes examined. *P* value for overall comparison by  $\chi^2$  test.

much of the global effect (Table 4).

Trying to dissect the association signal, we performed an analysis to control for individual markers one by one. We observed that no single tag SNP could explain the entire omnibus test results. By analyzing paired haplotypes, we observed that most of the paired haplotypes composed by tSNPs rs1554483, rs11932595, rs6850524 and rs4864548 were significantly associated with NAFLD (data not shown). Furthermore, the haplotype GAGA of these four tSNPs was significantly associated with NAFLD (OR: 2.13, 95% CI: 1.33-3.44, nominal *P* = 0.0011, empirical *P* = 0.00849).

## DISCUSSION

We examined the genetic influence of the gene variants and haplotypes of the LD block of the *CLOCK* transcription factor in a case-control study of NAFLD, and found that, in the analysis of individual markers, rs11932595 and rs6843722 were significantly associated with the disease being rs11932595A and rs6843722C allele carriers 1.6-fold more likely to have NAFLD in comparison with non-carriers. Moreover, we observed that *CLOCK* gene variant haplotypes frequencies in NAFLD individuals significantly differed from those in control subjects, indicating that either the GACCGA haplotype confers a 2-fold increase in the risk of suffering from NAFLD. Conversely, the CGCACG haplotype seems to confer a 3-fold protection against NAFLD, although the results from less frequent haplotypes should be taken with caution. By analyzing paired haplotypes, we observed that most of the paired haplotype composed by tSNPs rs1554483, rs11932595, rs6850524 and rs4864548 were significantly associated with NAFLD. Furthermore, the haplotype GAGA of these four tSNPs was significantly associated with NAFLD, indicating that subjects carrying this haplotype are more than twice as likely to have NAFLD in comparison with non-carriers.

Additionally, we evaluated the role of the gene variants in the disease severity and observed a significant association between the clinical or histological spectrum of NAFLD and the rs1554483, rs6843722 and rs6850524 tag SNPs.

Finally, the haplotype analysis revealed a significant association with the NAFLD, while no specific association

signal for any single tSNPs could explain this positive result. This observation seems to indicate the presence of un-typed causal variants associated with the disease.

To our knowledge, the potential contribution of *CLOCK* to the NAFLD susceptibility and severity has not been elucidated in humans and our study is probably the first to provide evidence of association to the disease. However, previous evidence in a mice model carrying a loss-of-function mutation in the *CLOCK* gene supports our findings as the animals develop adipocyte hypertrophy and excessive accumulation of fat in the liver<sup>[21]</sup>.

The *CLOCK* gene encodes a basic helix-loop-helix transcription factor that is essential for the circadian rhythm<sup>[50]</sup>. The circadian clock system located within the hypothalamic suprachiasmatic nucleus orchestrates the cascade of events in the central nervous system (CNS) that controls timing of food intake and metabolic processes as well as other physiological rhythms<sup>[51]</sup>. Circadian and metabolic processes are also regulated by *CLOCK* genes in peripheral tissues, particularly those involved in nutrient homeostasis such the liver<sup>[52]</sup>. Furthermore, in the mammalian circadian hierarchy, the liver may be primarily coupled to the CNS through rhythmic behaviour, such as feeding<sup>[52]</sup>.

Despite the fact that the molecular mechanisms underlying our observations are unknown, one possible explanation is that the gene variants in the *CLOCK* gene may lead to altered gene expression not only in the CNS but also in the liver.

Recently published findings about the role of gene variation on gene expression strengthen the hypothesis that variation in DNA sequence may contribute to individual differences in quantitative traits and disease susceptibility<sup>[33]</sup>. Interestingly, among 1097 genes and about 2 million SNPs tested for differences in expression phenotypes, *CLOCK* is listed as showing significant differences among ethnic groups, mostly owing to genetic variation in conserved non-coding regions<sup>[34]</sup>.

It is not clear why some persons develop simple steatosis, whereas others progress to severe cirrhosis. However, the fact that NASH is observed only in a fraction of patients with NAFLD clearly suggests a genetic predisposition. We observed that when we divided the NAFLD patients into categories of cases according to the clinical or histologic spectrum, a significant association was seen between the disease progression and the above-mentioned tSNPs (rs1554483, rs6843722 and rs6850524). Furthermore, tSNPs were significantly associated with the fibrosis score of the disease.

Although association does not necessarily mean a causal-relation, in this case, several lines of evidence support the relation between the *CLOCK* gene variants and NAFLD severity. For instance, a comprehensive and quantitative approach surveying the diurnal expression profiles of 49 mouse nuclear receptors (NRs) in white and brown adipose tissue, and liver and skeletal muscle revealed that 25 of 45 expressed NRs showed a rhythmic cycle<sup>[35]</sup>. Hence, the dynamic circadian changes in NRs expression may offer a logical explanation for our findings. Some of these NRs have already been reported to be involved in the pathogenesis of NASH (i.e. nuclear peroxisome

proliferator-activated receptors, PPARs) or in promoting liver fibrosis (i.e. retinoic acid and farnesoid X receptor)<sup>[36]</sup>. Not surprisingly, it was shown that the retinoid receptors can interact with *CLOCK* and an ortholog of *CLOCK* in a hormone-dependent manner<sup>[37]</sup>. Additionally, PPAR $\gamma$  is one of the metabolic transcription factors closely related to the coupled feedback pathway of the circadian oscillators<sup>[38]</sup>.

Lastly, activation of hepatic stellate cells (HSC) from lipid-storing pericytes to myofibroblastic cells constitutes a major cellular event in the genesis of liver fibrosis. It has recently been shown that putative adipogenic transcription factors are expressed in quiescent HSC. Many of these factors are under the control of PPAR $\gamma$  and sterol response element-binding proteins (SREBP-1)<sup>[39]</sup>. Interestingly, among the transcription factors of the *CLOCK* machinery, SREBP was shown to play an important role in the circadian organization of the liver transcriptome<sup>[40]</sup>.

In summary, genetic factors contribute to virtually every human disease by conferring susceptibility or resistance, affecting the severity or progression of disease, and interacting with environmental factors that modify disease course and expression. Our study suggests a potential role of the *CLOCK* gene variants and their related haplotypes in increased susceptibility to NAFLD and disease progression. Understanding how the reported variants located in different introns can affect the Clock function may require additional studies. However, it is worth mentioning that even robust genetic findings cannot always be readily explained at a functional level; however they are seeds for future research<sup>[41]</sup>.

We hope our study can serve as a primer since further research is needed to confirm and extend the current findings in a larger population so as to reveal the intimate mechanism by which the *CLOCK* variants may lead to the mentioned phenotype.

## COMMENTS

### Background

The pathogenesis of NAFLD is multifactorial and, as a complex disease, the disorder develops from the interplay between genes and environment. In mammals, many physiological processes are driven by a circadian timing system comprising a master pacemaker in the suprachiasmatic nuclei of the hypothalamus and peripheral oscillators in most body cells. Besides, many transcripts that participate in common metabolic pathways, such as metabolism of glucose, cholesterol and fatty acids, show circadian rhythmicity

### Research frontiers

The present study examined the genetic influence of the gene variants and haplotypes of the LD block of the *CLOCK* transcription factor in a case-control study of nonalcoholic fatty liver disease. The role of common polymorphisms in the abovementioned gene as indicators of disease risk is discussed.

### Innovations and breakthroughs

It is known that while the majority of patients with risk factors, such as obesity or insulin resistance, develop hepatic steatosis, only a minority progress to a more advanced liver disease (steatohepatitis or cirrhosis). Therefore, it is important to be able to identify individuals with fatty liver for intervention before this condition progresses. Recent advances in functional genomics are providing us great opportunities to decipher the interconnecting networks of genes, proteins, and metabolites. The use of candidate gene approach found *CLOCK* gene variants to be associated with an increased risk of NASH and also an important role in the progression of the disease.

## Applications

The elucidation of genetic risk factors for nonalcoholic fatty liver disease will allow development of susceptibility testing for disease prediction, particularly important for high-risk individuals.

## Terminology

Linkage disequilibrium (LD) is a term used in the study of population genetics for the non-random association of alleles at two or more loci, not necessarily on the same chromosome. Tag SNPs is a representative SNPs in a region of the genome with high linkage disequilibrium. It is possible to identify genetic variation without genotyping every SNP in a chromosomal region. Tagger is a tool for the selection and evaluation of tag SNPs from genotype data, such as that from the International HapMap Project. Aggressive tagging approach is searching for multimarker predictors as effective surrogates for single tag SNPs using the web-based service of the Tagger computer program.  $r^2$  is the minimal coefficient of determination at which all alleles are to be captured. WHAP program is a method for testing SNP haplotype associations with qualitative and quantitative traits in samples of individuals with or without parental genotype data. Basic-helix-loop-helix is a protein structural motif that characterizes a family of transcription factors.

## Peer review

This study is well conceived, well done and well presented.

## REFERENCES

- Desapriya E. Obesity epidemic. *Lancet* 2004; **364**: 1488
- Kopelman PG. Obesity as a medical problem. *Nature* 2000; **404**: 635-643
- Marchesini G, Bugianesi E, Forlani G, Cerrelli F, Lenzi M, Manini R, Natale S, Vanni E, Villanova N, Melchionda N, Rizzetto M. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *Hepatology* 2003; **37**: 917-923
- Bellentani S, Saccoccio G, Masutti F, Croce LS, Brandi G, Sasso F, Cristanini G, Tiribelli C. Prevalence of and risk factors for hepatic steatosis in Northern Italy. *Ann Intern Med* 2000; **132**: 112-117
- Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology* 1999; **116**: 1413-1419
- Neuschwander-Tetri BA, Caldwell SH. Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. *Hepatology* 2003; **37**: 1202-1219
- Chakravarti A, Little P. Nature, nurture and human disease. *Nature* 2003; **421**: 412-414
- Juran BD, Lazaridis KN. Genomics and complex liver disease: Challenges and opportunities. *Hepatology* 2006; **44**: 1380-1390
- Thomas EL, Brynes AE, Hamilton G, Patel N, Spong A, Goldin RD, Frost G, Bell JD, Taylor-Robinson SD. Effect of nutritional counselling on hepatic, muscle and adipose tissue fat content and distribution in non-alcoholic fatty liver disease. *World J Gastroenterol* 2006; **12**: 5813-5819
- Day CP. Genes or environment to determine alcoholic liver disease and non-alcoholic fatty liver disease. *Liver Int* 2006; **26**: 1021-1028
- Namikawa C, Shu-Ping Z, Vyselaar JR, Nozaki Y, Nemoto Y, Ono M, Akisawa N, Saibara T, Hiroi M, Enzan H, Onishi S. Polymorphisms of microsomal triglyceride transfer protein gene and manganese superoxide dismutase gene in non-alcoholic steatohepatitis. *J Hepatol* 2004; **40**: 781-786
- Merriman RB, Aouizerat BE, Bass NM. Genetic influences in nonalcoholic fatty liver disease. *J Clin Gastroenterol* 2006; **40**: S30-S33
- Piao YF, Li JT, Shi Y. Relationship between genetic polymorphism of cytochrome P450IIIE1 and fatty liver. *World J Gastroenterol* 2003; **9**: 2612-2615
- Song J, da Costa KA, Fischer LM, Kohlmeier M, Kwock L, Wang S, Zeisel SH. Polymorphism of the PEMT gene and susceptibility to nonalcoholic fatty liver disease (NAFLD). *FASEB J* 2005; **19**: 1266-1271
- Utzschneider KM, Kahn SE. Review: The role of insulin resistance in nonalcoholic fatty liver disease. *J Clin Endocrinol Metab* 2006; **91**: 4753-4761
- Hamaguchi M, Kojima T, Takeda N, Nakagawa T, Taniguchi H, Fujii K, Omatsu T, Nakajima T, Sarui H, Shimazaki M, Kato T, Okuda J, Ida K. The metabolic syndrome as a predictor of nonalcoholic fatty liver disease. *Ann Intern Med* 2005; **143**: 722-728
- Mendoza J. Circadian clocks: setting time by food. *J Neuroendocrinol* 2007; **19**: 127-137
- Wijnen H, Young MW. Interplay of circadian clocks and metabolic rhythms. *Annu Rev Genet* 2006; **40**: 409-448
- Panda S, Antoch MP, Miller BH, Su AI, Schook AB, Straume M, Schultz PG, Kay SA, Takahashi JS, Hogenesch JB. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* 2002; **109**: 307-320
- Delaunay F, Laudet V. Circadian clock and microarrays: mammalian genome gets rhythm. *Trends Genet* 2002; **18**: 595-597
- Turek FW, Joshi C, Kohsaka A, Lin E, Ivanova G, McDearmon E, Laposky A, Losee-Olson S, Easton A, Jensen DR, Eckel RH, Takahashi JS, Bass J. Obesity and metabolic syndrome in circadian Clock mutant mice. *Science* 2005; **308**: 1043-1045
- Mendler MH, Bouillet P, Le Sidaner A, Lavoine E, Labrousse F, Sautereau D, Pillegand B. Dual-energy CT in the diagnosis and quantification of fatty liver: limited clinical value in comparison to ultrasound scan and single-energy CT, with special reference to iron overload. *J Hepatol* 1998; **28**: 785-794
- Haffner SM, Kennedy E, Gonzalez C, Stern MP, Miettinen H. A prospective analysis of the HOMA model. The Mexico City Diabetes Study. *Diabetes Care* 1996; **19**: 1138-1141
- Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* 1999; **94**: 2467-2474
- Kawasaki ES. Sample preparation from blood, cells, and other fluids. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR Protocols: A Guide to Methods and Application. San Diego: Academic Press, 1990: 146-152
- de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet* 2005; **37**: 1217-1223
- Myakishev MV, Khripin Y, Hu S, Hamer DH. High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. *Genome Res* 2001; **11**: 163-169
- Purcell S, Daly MJ, Sham PC. WHAP: haplotype-based association analysis. *Bioinformatics* 2007; **23**: 255-256
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263-265
- King DP, Zhao Y, Sangoram AM, Wilsbacher LD, Tanaka M, Antoch MP, Steeves TD, Vitaterna MH, Kornhauser JM, Lowrey PL, Turek FW, Takahashi JS. Positional cloning of the mouse circadian clock gene. *Cell* 1997; **89**: 641-653
- Reppert SM, Weaver DR. Coordination of circadian timing in mammals. *Nature* 2002; **418**: 935-941
- Stokkan KA, Yamazaki S, Tei H, Sakaki Y, Menaker M. Entrainment of the circadian clock in the liver by feeding. *Science* 2001; **291**: 490-493
- Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavare S, Deloukas P, Hurles ME, Dermitzakis ET. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 2007; **315**: 848-853
- Spiegel K, Knutson K, Leproult R, Tasali E, Van Cauter E. Sleep loss: a novel risk factor for insulin resistance and Type 2 diabetes. *J Appl Physiol* 2005; **99**: 2008-2019
- Yang X, Downes M, Yu RT, Bookout AL, He W, Straume M, Mangelsdorf DJ, Evans RM. Nuclear receptor expression links the circadian clock to metabolism. *Cell* 2006; **126**: 801-810
- Gaemers IC, Groen AK. New insights in the pathogenesis of non-alcoholic fatty liver disease. *Curr Opin Lipidol* 2006; **17**: 268-273

- 37 **McNamara P**, Seo SP, Rudic RD, Sehgal A, Chakravarti D, FitzGerald GA. Regulation of CLOCK and MOP4 by nuclear hormone receptors in the vasculature: a humoral mechanism to reset a peripheral clock. *Cell* 2001; **105**: 877-889
- 38 **Staels B**. When the Clock stops ticking, metabolic syndrome explodes. *Nat Med* 2006; **12**: 54-55; discussion 55
- 39 **She H**, Xiong S, Hazra S, Tsukamoto H. Adipogenic transcriptional regulation of hepatic stellate cells. *J Biol Chem* 2005; **280**: 4959-4967
- 40 **Brewer M**, Lange D, Baler R, Anzulovich A. SREBP-1 as a transcriptional integrator of circadian and nutritional cues in the liver. *J Biol Rhythms* 2005; **20**: 195-205
- 41 **Craddock N**, Owen MJ, O'Donovan MC. The catechol-O-methyl transferase (COMT) gene as a candidate for psychiatric phenotypes: evidence and lessons. *Mol Psychiatry* 2006; **11**: 446-458

S- Editor Liu Y L- Editor Kumar M E- Editor Liu Y