Chapter 9
Chronobiology and Metabolic Syndrome: From Genes to Systems Biology

Silvia Sookoian and Carlos J. Pirola

Abstract The major function of the circadian system is the internal cycling of physiologic and metabolic events. In the last years, there has been an exponential increase in our understanding of the role of clock-related genes in Metabolic-syndrome (MetS)-related phenotypes. Nevertheless, our understanding about how the components of the circadian system interact each other to modulate the metabolism and cardiovascular system remains a major challenge.

Systems biology introduces a new concept for revealing the pathogenesis of human disorders and suggests the presence of common physiologic processes and molecular networks influencing the risk of a disease. In this review, we use systems biology approaches to integrate genomic, molecular, and physiological data to decipher putative circadian rhythmic pathways suspected to play a role in the etiology of the metabolic syndrome (MetS)-associated phenotypes with a main focus in obesity and as other authors are discussing diverse related topics, we discuss the findings of our own group.

S. Sookoian, M.D., Ph.D. (✉)
Instituto de Investigaciones Médicas A. Lanari,
Av. Combatiente de Malvinas 3150, (1427) Buenos Aires, Argentina

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), Ciudad Autónoma de Buenos Aires, Argentina
e-mail: sookoian.silvia@lanari.fmed.uba.ar

C.J. Pirola, Ph.D. (✉)
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), Ciudad Autónoma de Buenos Aires, Argentina
e-mail: pirola.carlos@lanari.fmed.uba.ar

M. Garaulet and J.M. Ordovás (eds.), *Chronobiology and Obesity*,
DOI 10.1007/978-1-4614-5082-5_9,

Abbreviations

MetS	Metabolic syndrome
IR	Insulin resistance
T2D	Type-2 diabetes
GWAS	Genome-wide association scan
CLOCK	Circadian locomotor output cycles kaput protein
5-HT	Serotonin
5-HTT	Serotonin transporter
5HIAA	5-Hydroxyindolacetic acid
PPARG1A	Peroxisome proliferator-activated receptor gamma cofactor 1 alpha gene

Introduction

The Impact of the Circadian System in the Metabolic Syndrome-Associated Phenotypes

Metabolic syndrome (MetS) is associated with several metabolic disturbances, including insulin resistance (IR) in several tissues. Indeed, IR is considered as the main link among all the clinical disorders clustered in MetS, namely type-2 Diabetes (T2D), dyslipidemias, central obesity, arterial hypertension, prothrombotic and proinflammatory states, ovarian polycystosis, and nonalcoholic fatty liver disease (NAFLD). From the perspective of clinical importance, MetS has two prominent features: Its worldwide prevalence that is dramatically increasing and its strong association with cardiovascular disease, as initially described by Reaven G [1].

The pathobiology of the MetS results from a complex interplay between genes and environment. Actually, environmental factors, such as decreased physical activity, increased nutrient availability, and overnutrition, play an important role in the development of metabolic disturbances associated with IR, and are also largely recognized as responsible for the modern epidemic of MetS-related phenotypes.

The genetic component of each individual component of the MetS has been largely investigated and both genome-wide and candidate gene association studies have identified several loci that influence the susceptibility of all the clustering traits [2, 3].

Nevertheless, the gene variants identified so far by genome-wide association scans (GWAS) do not explain by themselves the pathophysiology of the MetS, and they account for a modest effect on the disease. There has been a loci that was not initially identified by any of the GWAS but showed an interesting biological plausibility and a significant effect on some intermediate phenotypes associated with MetS, such as obesity. This locus is related to the circadian system, and is actually the master gene circadian locomotor output cycles kaput protein (*CLOCK*).

Actually, the importance of maintaining the internal homeostasis of the circadian systems and its impact on human MetS was recently revealed by Turek et al. by an animal model in which homozygous *Clock* mutant mice had a greatly attenuated diurnal feeding rhythm, were hyperphagic and obese, and developed hyperleptinemia, hyperlipidemia, hepatic steatosis, hyperglycemia, and hypoinsulinemia [4]. This finding is biologically plausible as the major function of the circadian system is the internal cycling of physiologic and metabolic events [5].

Given the above mentioned evidence and the results of further emerging studies, which show that altering circadian rhythmicity results in pathophysiological changes resembling MetS and fat accumulation, we immediately explored the role of gene variants and derived haplotypes of the *CLOCK* transcription factor in obesity and related quantitative metabolic traits in a human study [6]. Hence, we recruited in a case–control design, 715 lean and 391 overweight or obese unrelated adult subjects, and investigated six tag single-nucleotide polymorphisms (SNPs) with 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. The results of our study suggested that *CLOCK* polymorphism and related haplotypes are critically involved in the genetic susceptibility to obesity as carrying the haplotype of rs1554483G and rs4864548A was associated with 1.8-fold increase for being overweight/obese [6]. These findings were replicated in other populations around the word [7–9]. In addition, although the association of the *CLOCK* variants with other classical components of the MetS such as arterial hypertension did not persist after adjusting for overweight/obesity, we did find a significant association with non alcoholic fatty liver disease (NAFLD) in a hospital-based study [10]. In fact, the *CLOCK* variants, rs11932595 and rs6843722 showed significant associations with NAFLD (empiric $P = 0.0449$ and 0.023, respectively), and a significant association was also observed between clinical or histologic spectrum of the disease 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 $P = 0.02697$) suggesting a potential role of the *CLOCK* polymorphisms and their haplotypes in the susceptibility to NAFLD and disease severity [10]. A more comprehensive review of the genetic basis of the MetS was reported elsewhere [3].

In summary, while in the last years there has been an exponential increase in our understanding of the role of clock-related genes in MetS-related phenotypes [11], to know how the components of the circadian system interact with each other to modulate the metabolism and cardiovascular system remains a major challenge.

A Short Overview About Circadian System–Gene Components

The circadian system of mammals is composed of a hierarchy of oscillators that function at the cellular, tissue, and systems levels (Fig. 9.1) [12]. Component of the

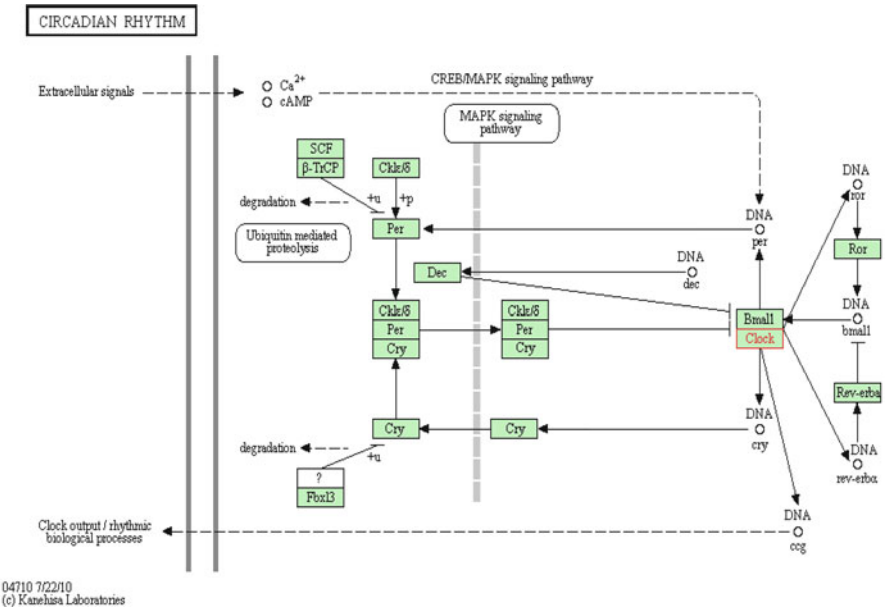


Fig. 9.1 Circadian rhythm system component of mammal—*Homo sapiens* (human) according to the KeGG (Kyoto Encyclopedia of Genes and genomes) map 04710

circadian clock oscillator includes the CRY proteins, CLOCK or NPAS2, ARNTL or ARNTL2, CSNK1D and/or CSNK1E, TIMELESS and the PER proteins. Table 9.1 summarizes the main features of the transcription/translation feedback loops of the mammalian circadian clock. As this picture is an oversimplification, here, we decide to explore the association of obesity and circadian rhythm-related genes using systems biology.

The Use of a Systems Biology Approach to Better Understand the Association of Circadian Rhythm-Related Genes with Obesity and MetS

Systems biology introduces a new concept for revealing the pathogenesis of human disorders and suggests the presence of common physiologic processes and molecular networks influencing the risk of a disease. Rather than compartmentalizing individual risk factors (e.g., IR, blood pressure, body mass index, or lipid concentrations) and treating them as if they were separate and independent, systems biology examines their interactions. Here we show a model of this concept to explain the impact of the circadian system on one of the most important features of the MetS, obesity, by different systems-biology approaches, which are mostly based on gene enrichment analysis and protein–protein interaction networks.

Table 9.1 Short overview about the main function and features of the genes related with the circadian system

Gene symbol	Gene name	Main function and features
CLOCK	Circadian locomotor output cycles kaput protein	• Protein involved in the generation of rhythmic pattern of behaviors or activities, e.g., circadian rhythm which is a metabolic or behavioral rhythm within a cycle of 24 h. • ARNTL/2-CLOCK heterodimers activate E-box element transcription of a number of proteins of the circadian clock. • Activates transcription of PER1 and PER2. • Has intrinsic histone acetyltransferase activity and this enzymatic function contributes to chromatin-remodeling events implicated in circadian control of gene expression. • CLOCK gene, homolog to murine Clock, regulator of circadian rhythms, with two major transcripts 8 and 10 kb, predominantly expressed in the suprachiasmatic nuclei and cerebellum. • Expressed in all peripheral tissues, including spleen, thymus, prostate, testis, ovary, small intestine, colon, leukocytes, heart, brain, placenta, lung, liver, skeletal muscle, kidney, adipose tissue and pancreas.
CRY1	Cryptochrome 1 (photolyase-like)	• Blue light-dependent regulator of the circadian feedback loop. Inhibits CLOCK NPAS2-ARNTL E box-mediated transcription. • Acts, in conjunction with CRY2, in maintaining period length and circadian rhythmicity. • Capable of translocating circadian clock core proteins such as PER proteins to the nucleus. • Expression is regulated by light and circadian rhythms. Peak expression in the suprachiasmatic nucleus (SCN) and eye at the day/night transition (CT12). Levels decrease with ARNTL-CLOCK inhibition as part of the autoregulatory feedback loop. • Belongs to the DNA photolyase class-1 family.
CRY2	Cryptochrome 2 (photolyase-like)	• Blue light-dependent regulator of the circadian feedback loop • Translocated to the nucleus through interaction with other Clock proteins such as PER2 or ARNTL • Belongs to the DNA photolyase class-1 family. • Regulate the functional activity of circadian transcriptional complex at the posttranslational level • Phosphorylation on Ser-266 by MAPK is important for the inhibition of CLOCK-ARNTL-mediated transcriptional activity. Phosphorylation by CSKNE requires interaction with PER1 or PER2. • Expressed in all tissues examined including fetal brain, fibroblasts, heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon, and leukocytes. Highest levels in heart and skeletal muscle.

(continued)

Table 9.1 (continued)

Gene symbol	Gene name	Main function and features
PER1/PER2	Period homolog 1 (Drosophila) and period homolog 2 (Drosophila)	• Component of the circadian clock mechanism which is essential for generating circadian rhythms. • Negative element in the circadian transcriptional loop. Influences clock function by interacting with other circadian regulatory proteins and transporting them to the nucleus. • Negatively regulates CLOCK NPAS2-BMAL1 BMAL2-induced transactivation. • PER1: Widely expressed. Found in heart, brain, placenta, lung, liver, skeletal muscle, pancreas, kidney, spleen, thymus, prostate, testis, ovary, and small intestine. Highest level in skeletal muscle. Low level in kidney. • PER2: Widely expressed. Found in heart, brain, placenta, lung, liver, skeletal muscle, kidney, adipose tissue and pancreas. High levels in skeletal muscle and pancreas. Low level in lung.
ARNTL	Aryl hydrocarbon receptor nuclear translocator-like (alias BMAL)	• The protein encoded by this gene is a basic helix-loop-helix protein that forms a heterodimer with CLOCK. • This complex binds an E-box upstream of the PER1 gene, activating this gene and possibly other circadian rhythm-associated genes. • master regulator of circadian rhythm, also playing important roles in the regulation of adipose differentiation and lipogenesis in mature adipocytes.
CSNK1D	Casein kinase 1, delta	• This gene is a member of the casein kinase I (CKI) gene family whose members have been implicated in the control of cytoplasmic and nuclear processes, including DNA replication and repair. • Phosphorylates connexin-43/GJA1, MAP1A, SNAPIN, MAPT/TAU, TOP2A, DCK, HIF1A, EIF6, p53/TP53, DVL2, DVL3, ESR1, AIB1/NCOA3, DNMT1, PKD2, YAP1, PER1, and PER2. • Central component of the circadian clock. May act as a negative regulator of circadian rhythmicity by phosphorylating PER1 and PER2, leading to retain PER1 in the cytoplasm.
TIMELESS	*Timeless* homolog (Drosophila)	• Required for normal progression of S-phase. • Involved in the circadian rhythm autoregulatory loop. • Negatively regulates CLOCK-NPAS2/BMAL1-induced transactivation of PER1 possibly via translocation of PER1 into the nucleus. • Expressed in all tissues examined including brain, heart, lung, liver, skeletal muscle, kidney, placenta, pancreas, spleen, thymus, and testis. Highest levels of expression in placenta, pancreas, thymus and testis. • Belongs to the timeless family.

For example, we used a bioinformatic resource named "Platform for Exploration of Significant Concepts Associated to co-Ocurrences Relationships" (Pescador, cbdm.mdc-berlin/tools/pescador) [13].

PESCADOR uses LAITOR-Literature Assistant for Identification of Terms co-Occurrences and Relationships as text-mining engine to extract sentences with co-occurring bioentities (genes and proteins) from the text of the PubMed abstracts requested [13]. Thus, PESCADOR allows selecting gene–protein co-occurrence pairs based on their relatedness to biological concepts and therefore brings together under a common perspective protein interactions that have not been studied under the same research focus [13].

Thus, we used the keywords, "obesity AND genetics AND circadian" retrieving 200 references. The associated genes under this search are described in Table 9.2. Interestingly, more than 100 genes associated with the regulation of food intake and energy expenditure, glucose and lipid metabolism, and of course circadian rhythmicity, as displayed in Fig. 9.2, participate in a highly interconnected interactome. Interestingly, the hierarchical order of the genes shows that *CLOCK* is the most important one, followed by other components of the cell circadian clock mechanism (Fig. 9.3). The notion of "interactome" includes the complete list of physical interactions mediated by all proteins of an organism [14]. Indeed, biological questions are increasingly addressed in the framework of such complex molecular networks (Tables 9.3 and 9.4).

Based on the hypothesis that common physiologic processes and molecular networks influence the risk of obesity, we proposed another systems biology approach: a gene enrichment analysis evaluated by the bioinformatic resource *ToppGene Suite* (http://toppgene.cchmc.org) [15]. A similar concept was recently applied for finding new candidate genes for the MetS [2], rendering new loci whose associations with MetS components were finally replicated in independent studies, i.e., *HNF4A* with type 2 diabetes in more than 49,000 individuals by a meta-analysis [16], and *IGF1R* with arterial hypertension in an Argentinean population [17]. Importantly, by using the ToppFun (Transcriptome, ontology, phenotype, proteome, and pharmacome annotations based gene list functional enrichment analysis, a tool of the Toppgene suite), is possible to obtain a general picture of the biological process these genes are belonging to with a significant P value $< 1 \times 10^{-6}$. In this sense, it is important to note that in addition to the obvious pathways already mentioned, an important number of loci, i.e., 25%, belong to biological process involved in regulation of the transcriptional activity of polymerase II probably as nuclear factors "per se" or through their activity on epigenetic marks, i.e., histone deacetylases (HDACs, SIRT1, etc.). This topic is further developed below.

Some other genes are involved in the response to xenobiotics (according to ToppFun, P value $< 1 \times 10^{-6}$), and then it is worth of noting that not only are some clinically relevant drugs but natural compound mentioned. In this regards, it is particularly interesting that several serotonin transporter inhibitors, such as sibutramine or fenfluramine and serotonin itself are included.

Table 9.2 List of genes retrieved by the Pescador tool (available at http://cbdm.mdc-berlin.de/tools/pescador/)

Gene symbol	Gene name
ACACB	Acetyl-CoA carboxylase beta
ACSS2	Acyl-CoA synthetase short-chain family member 2
ADIPOQ	Adiponectin, C1Q and collagen domain containing
ADIPOR1	Adiponectin receptor 1
ADIPOR2	adiponectin receptor 2
AGRP	Agouti related protein homolog (mouse)
ARNTL	Aryl hydrocarbon receptor nuclear translocator-like
ATP2A2	ATPase, Ca2+ transporting, cardiac muscle, slow twitch 2
BDNF	Brain-derived neurotrophic factor
BTG2	BTG family, member 2
C2	Complement component 2
CCRN4L	CCR4 carbon catabolite repression 4-like (*S. cerevisiae*)
CD36	CD36 molecule (thrombospondin receptor)
CLOCK	Clock homolog (mouse)
COL4A5	Collagen, type IV, alpha 5
CREM	cAMP responsive element modulator
CRY1	Cryptochrome 1 (photolyase-like)
CRY2	Cryptochrome 2 (photolyase-like)
CYP19A1	Cytochrome P450, family 19, subfamily A, polypeptide 1
DBP	D site of albumin promoter (albumin D-box) binding protein
ETV3	Ets variant 3
FAS	Fas (TNF receptor superfamily, member 6)
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GCG	Glucagon
GCGR	Glucagon receptor
GCK	Glucokinase (hexokinase 4)
GHRH	Growth hormone releasing hormone
GHRL	Ghrelin/obestatin prepropeptide
GLB1	Galactosidase, beta 1
GLP1R	Glucagon-like peptide 1 receptor
GLP2R	Glucagon-like peptide 2 receptor
GPR50	G protein-coupled receptor 50
GPT2	Glutamic pyruvate transaminase (alanine aminotransferase) 2
H6PD	Hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)
HCRT	Hypocretin (orexin) neuropeptide precursor
HDAC1	Histone deacetylase 1
HDAC3	Histone deacetylase 3
HR	Hairless homolog (mouse)
ID1	Inhibitor of DNA binding 1, dominant negative helix–loop–helix protein
ID2	Inhibitor of DNA binding 2, dominant negative helix–loop–helix protein
INS	Insulin
LEP	Leptin
LEPR	Leptin receptor
LTF	Lactotransferrin
MC3R	Melanocortin 3 receptor

(continued)

Table 9.2 (continued)

Gene symbol	Gene name
MC4R	Melanocortin 4 receptor
MECP2	Methyl CpG binding protein 2 (Rett syndrome)
MT1E	Metallothionein 1E
MT2A	Metallothionein 2A
MTNR1A	Melatonin receptor 1A
MTNR1B	Melatonin receptor 1B
NAMPT	Nicotinamide phosphoribosyltransferase
NCOR1	Nuclear receptor corepressor 1
NFIL3	Nuclear Factor, interleukin 3 regulated
NOVA2	Neuro-oncological ventral antigen 2
NPAS2	Neuronal PAS domain protein 2
NPY	Neuropeptide Y
NR1I2	Nuclear receptor subfamily 1, group I, member 2
NR1I3	Nuclear receptor subfamily 1, group I, member 3
NR3C1	Nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)
NUCB2	Nucleobindin 2
O3FAR1	Omega-3 fatty acid receptor 1
PAIP2	Poly(A) binding protein interacting protein 2
PCK2	Phosphoenolpyruvate carboxykinase 2 (mitochondrial)
PDK4	Pyruvate dehydrogenase kinase, isozyme 4
PER1	Period homolog 1 (Drosophila)
PER2	Period homolog 2 (Drosophila)
PER3	Period homolog 3 (Drosophila)
POMC	Proopiomelanocortin
PPA1	Pyrophosphatase (inorganic) 1
PPARA	Peroxisome proliferator-activated receptor alpha
PPARG	Peroxisome proliferator-activated receptor gamma
PRKAA1	Protein kinase, AMP-activated, alpha 1 catalytic subunit
PRL	Prolactin
PROK2	Prokineticin 2
PROKR2	Prokineticin receptor 2
PYY	Peptide YY
SDC3	Syndecan 3
SDC3 s	Yndecan 3
SERPINE1	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
SLC2A4	Solute carrier family 2 (facilitated glucose transporter), member 4
SLC5A1	Solute carrier family 5 (sodium/glucose cotransporter), member 1
SLC5A1	Solute carrier family 5 (sodium/glucose cotransporter), member 1
SLC9A3	Solute carrier family 9 (sodium/hydrogen exchanger), member 3
STAT3	Signal transducer and activator of transcription 3 (acute-phase response factor)
UBE4A	Ubiquitination factor E4A
UCP1	Uncoupling protein 1 (mitochondrial, proton carrier)
UCP3	Uncoupling protein 3 (mitochondrial, proton carrier)
VGF	VGF nerve growth factor inducible
YARS	Tyrosyl-tRNA synthetase

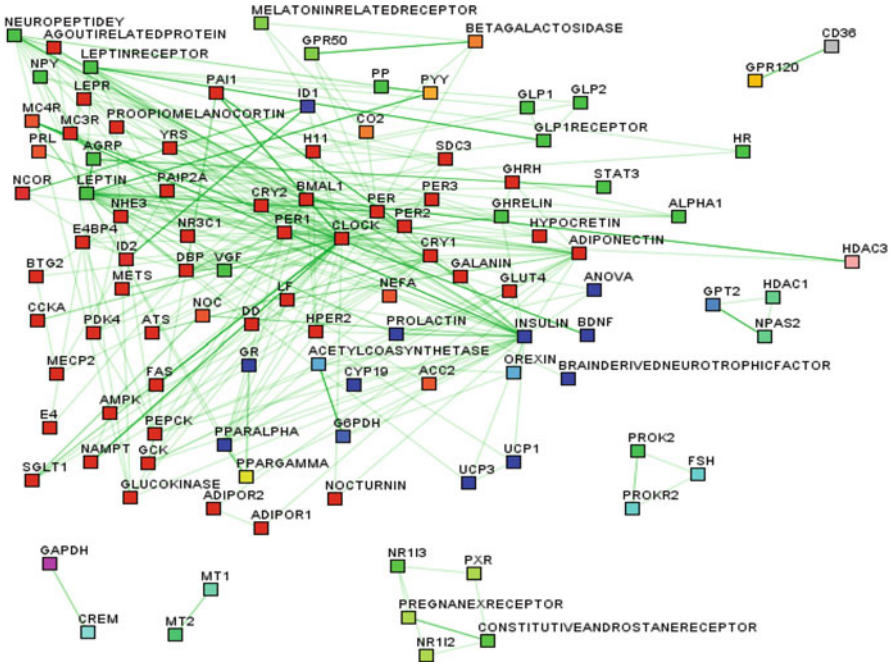


Fig. 9.2 Interactome of genes listed in Table 9.2

The Connection Between the Circadian System, Serotonin and Obesity

As an example of the application of systems biology to understand the pathobiology of human diseases, we explain in detail the results of a highly predicted protein by the above mentioned strategy: the serotonin transporter and its association with obesity.

Serotonin transporter (5-HTT) is involved in mood and eating disturbances and encoded by the gene *SLC6A4*, the promoter shows functional insertion/deletion alleles: long (L) and short (S). Because individuals who are carriers for the short version are known to be at risk for higher levels of anxiety, we hypothesized that this variant may be associated with overweight [18, 19]. We collected data and blood samples from 172 adolescents out of a cross-sectional, population-based study of 934 high school students and to replicate the findings, we also included 119 outpatients from the Nutrition and Diabetes Section of the Children's County Hospital [18]. We found that the S allele was associated with overweight (BMI > 85th percentile), being a risk factor for overweight independently of sex, age, and hypertension [odds ratio (OR): 1.85; 95% confidence interval (CI): 1.13, 3.05; $P < 0.02$] [18]. Additionally, in the outpatient study, compared with the homozygous LL subjects, S allele carriers showed a higher BMI z-score (1.47 ± 1.09 vs. 0.51 ± 1.4; $P < 0.002$)

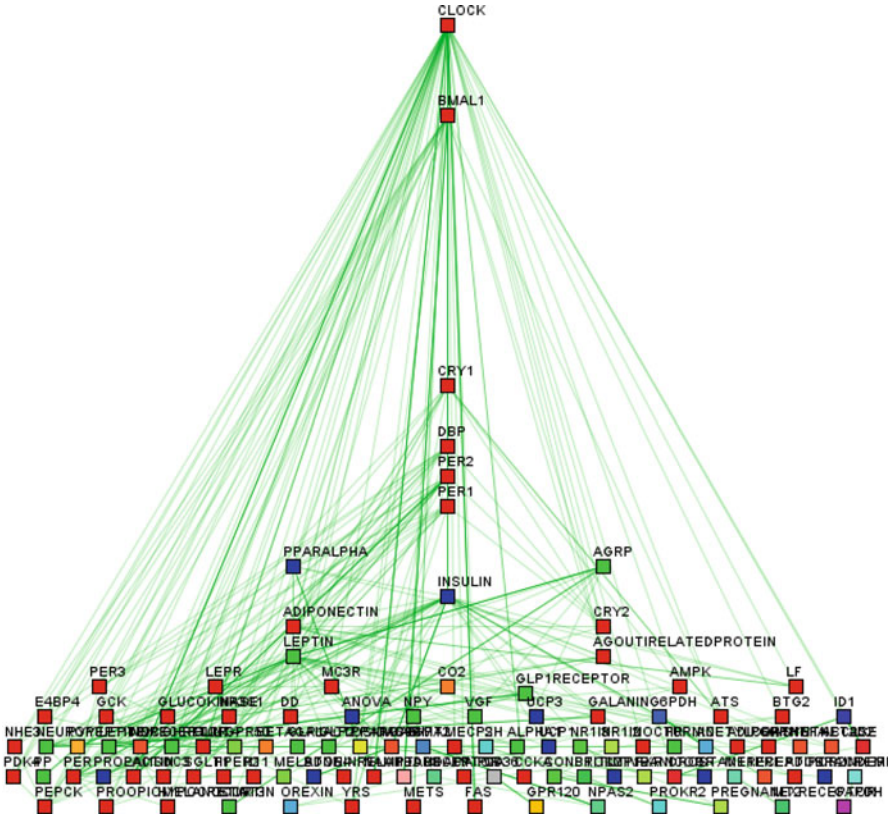


Fig. 9.3 Hierarchical interactome of genes listed in Table 9.2

and were more frequent in overweight children [18]. Furthermore, we further evaluated whether the S/L variant of the *SLC6A4* gene is associated with BMI as a continuous trait and also with obesity in a large sample of adult men of European ancestry included in a cross-sectional, population-based study (1,329 unrelated subjects, aged 34.6 ± 0.3 years) [19]. We observed statistically significant differences across genotypic groups (LL: 25.4 ± 0.2, LS: 26.0 ± 0.1 and SS: 26.7 ± 0.2, $P < 0.0002$). In addition, association tests showed that the 5-HTTLPR-genotype distribution was significantly different between 692 lean (BMI ≤ 25 kg/m^2) and 637 overweight/obese (BMI ≥ 27 kg/m^2) individuals, and we found a 1.36 odds ratio (OR) (95% CI 1.01–1.85) for obesity in SS carriers in comparison with LL carriers, $P = 0.026$ [19].Our findings indicate that 5-HTTLPR polymorphism may be linked with BMI and also with obesity and/or overweight in adolescent and adult populations reinforcing the role of the serotonin transporter as a risk factor for the obesity phenotype and suggest potential new avenues for its pharmacological treatment.

In fact, because serotonin (5-HT) is a neurotransmitter associated with circadian rhythm regulation, we explored a possible relation among 5-HT, its metabolite,

Table 9.3 Biological process to which listed genes are involved with as classified by Gene Ontology (GO) terms

ID of the GO term	Name of the GO term	Hit count in query list	Hit count in genome
GO:0007623	Circadian rhythm	21	85
GO:0048511	Rhythmic process	25	201
GO:0009725	Response to hormone stimulus	29	644
GO:0009719	Response to endogenous stimulus	29	714
GO:0010033	Response to organic substance	35	1,206
GO:0032870	Cellular response to hormone stimulus	19	345
GO:0043434	Response to peptide hormone stimulus	19	353
GO:0071495	Cellular response to endogenous stimulus	19	363
GO:0007631	Feeding behavior	12	94
GO:0006091	Generation of precursor metabolites and energy	19	443
GO:0070887	Cellular response to chemical stimulus	24	819
GO:0009605	Response to external stimulus	27	1,123
GO:0071310	Cellular response to organic substance	19	492
GO:0019318	Hexose metabolic process	15	261
GO:0031327	Negative regulation of cellular biosynthetic process	23	796
GO:0006006	Glucose metabolic process	14	217
GO:0009755	Hormone-mediated signaling pathway	11	102
GO:0009890	Negative regulation of biosynthetic process	23	810
GO:0032787	Monocarboxylic acid metabolic process	18	445
GO:0048878	Chemical homeostasis	22	740
GO:0055114	Oxidation–reduction process	24	935
GO:0032094	Response to food	7	22
GO:0007610	Behavior	18	477
GO:0031667	Response to nutrient levels	15	297
GO:0005996	Monosaccharide metabolic process	15	299
GO:0032102	Negative regulation of response to external stimulus	10	87
GO:0048545	Response to steroid hormone stimulus	15	303
GO:0032098	Regulation of appetite	6	13
GO:0032095	Regulation of response to food	6	13
GO:0042592	Homeostatic process	25	1,086
GO:0006366	Transcription from RNA polymerase II promoter	26	1,206
GO:0006357	Regulation of transcription from RNA polymerase II promoter	24	1,015
GO:0009892	Negative regulation of metabolic process	25	1,112
GO:0009991	Response to extracellular stimulus	15	325
GO:0010817	Regulation of hormone levels	16	386
GO:0031324	Negative regulation of cellular metabolic process	24	1,023
GO:0048585	Negative regulation of response to stimulus	12	174
GO:0010558	Negative regulation of macromolecule biosynthetic process	21	773
GO:0043436	Oxoacid metabolic process	22	863
GO:0019752	Carboxylic acid metabolic process	22	863
GO:0042593	Glucose homeostasis	9	73
GO:0033500	Carbohydrate homeostasis	9	73

(continued)

Table 9.3 (continued)

ID of the GO term	Name of the GO term	Hit count in query list	Hit count in genome
GO:0006082	Organic acid metabolic process	22	879
GO:0042180	Cellular ketone metabolic process	22	880
GO:0040014	Regulation of multicellular organism growth	9	78
GO:0032868	Response to insulin stimulus	13	249
GO:0032107	Regulation of response to nutrient levels	6	18
GO:0032104	Regulation of response to extracellular stimulus	6	18
GO:0006066	Alcohol metabolic process	18	580
GO:0046879	Hormone secretion	12	207

The Gene Ontology project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases. The project provides a controlled vocabulary of terms for describing gene product characteristics and gene product annotation data from GO Consortium members, as well as tools to access and process this data (http://www.geneontology.org/)

5-hydroxyindolacetic acid (5HIAA), and the functional polymorphism of the serotonin transporter gene (SLC6A4) promoter with rotating shift work. We decided to explore the impact of serotonin on this clinical model, as rotating shift work disrupts the synchronous relationship between the body's internal clock and the environment, causing deleterious effects not only on hormonal levels but also on neurotransmitters and activity–rest cycles. Thus, we performed a cross sectional study, including 683 men, and 437 day workers were compared with 246 rotating shift workers, and we found that platelet 5-HT content differed significantly ($P=0.002$) between day workers (41.28 ± 1.99 pg/mg) and rotating shift workers (37.91 ± -4.16 pg/mg); 5-HIAA content was also significantly ($P=0.00004$) higher in day workers (11.40 ± 0.82 pg/mg) than in rotating shift workers (9.33 ± 1.02 pg/mg). When we looked for differences in *SLC6A4* promoter, we found a significant ($P=0.016$) difference in genotype distribution between day workers LL: 126 (28.8%), LS: 202 (46.2%), and SS: 109 (24.9%), and rotating shift workers LL: 47 (19.1%), LS: 124 (50.4%), and SS: 75 (30.5%). Interestingly, when we divided the subjects between workers with less and more than 60 month rotating shift-work exposure, the difference in *SLC6A4* genotypes frequency was only significant in the group with > or =60 months ($P=0.011$). In addition, there was a significantly lower content of platelet 5-HIAA in S allele carriers in comparison with the other genotypes (SS: 9.2 ± 1.0 pg/mg vs. SL/LL: 11.0 ± 0.8 pg/mg, $P<0.02$). As a conclusion, platelet 5-HT and 5-HIAA contents were significantly lower in rotating shift workers than day workers, and there was a significant association between the S variant of SLC6A4 promoter and shift work. These findings may be important for targeting effective therapeutic strategies to ameliorate the associated comorbidities and behavioral problems in rotating shift workers, because, as it is well-known, rotating shift workers present many of the features of the MetS including low-grade inflammation and elevated circulating cytokines [20–22].

Table 9.4 Interacting xenobiotics with the listed genes

ID	Name	Source	Hit count in query list	Hit count in genome
CID000004829	Pioglitazone	Stitch	25	326
CID000000896	Melatonin	Stitch	26	443
CID000005753	Corticosterone	Stitch	24	381
CID000003003	Dexamethasone	Stitch	38	1,343
D005632	Fructose	CTD	16	125
CID000000861	Triiodothyronine	Stitch	23	409
CID000004599	Orlistat	Stitch	14	86
CID000003413	Forskolin	Stitch	32	1,014
D003907	Dexamethasone	CTD	26	624
CID000000274	Cyclic AMP	Stitch	30	907
MESH:D009765/ D005632-M	Obesity affected by fructose	CTD Marker	8	11
CID000005591	Troglitazone	Stitch	24	526
CID000004091	Metformin	Stitch	18	235
CID000005300	Streptozotocin	Stitch	24	540
CID000001106	Stearoyl-coenzyme A	Stitch	14	122
CID000102191	2-Deoxyglucose	Stitch	18	272
CID000000869	Malonyl-COA	Stitch	14	126
CID000005210	Sibutramine	Stitch	10	40
D005947	Glucose	CTD	14	143
CID000002900	Cycloheximide	Stitch	27	906
CID000005244	Dehydroepiandrosterone sulfate	Stitch	17	277
CID000005694	Wy-14,643	Stitch	17	287
CID000060303	Englitazone	Stitch	8	22
CID000077999	Rosiglitazone	Stitch	22	605
CID000145068	Nitric oxide	Stitch	28	1,076
CID000003339	Fenofibrate	Stitch	16	276

MESH:D009765/ D004041-M	Obesity affected by Dietary Fats	CTD Marker	8	27
CID000031275	1,4-Dioxane	Stitch	16	278
CID000003640	Cortisol	Stitch	20	504
CID000005408	Testosterone	Stitch	21	578
D010126	Ozone	CTD	25	882
CID000004920	Progesterone	Stitch	28	1,171
CID000004585	Olanzapine	Stitch	12	129
C023036	Perfluorooctanoic acid	CTD	20	543
CID000005202	Serotonin	Stitch	20	554
CID000005098	GI262570	Stitch	11	104
D005227	Fatty Acids	CTD	8	34
CID000000304	Cholesterol	Stitch	21	637
CID000003337	Fenfluramine	Stitch	13	178
CID000004174	Metyrapone	Stitch	13	179
CID000000681	Dopamine	Stitch	22	715
CID000002826	Cocaine	Stitch	20	571
CID000002833	Colchicine	Stitch	17	383
C055122	Orlistat	CTD	9	56
CID000003285	Estrogen	Stitch	28	1,252
C009250	Sevoflurane	CTD	12	146
CID000003779	Isoproterenol	Stitch	21	669
CID000000702	EtOH	Stitch	23	837
CID000000076	Dehydroepiandrosterone	Stitch	15	292
D019324	Beta-naphthoflavone	CTD	20	608
CID000036523	Synthetic LH-RH	Stitch	14	246
CID000000200	5-Aminoimidazole-4-carboxamide ribotide	Stitch	12	163

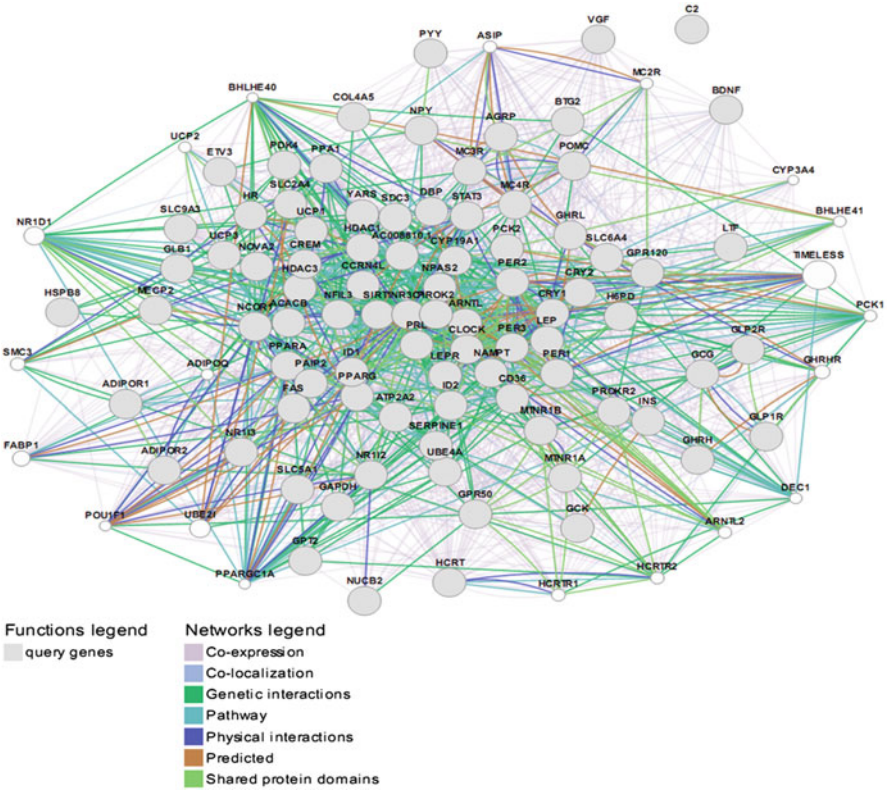
(continued)

Table 9.4 (continued)

ID	Name	Source	Hit count in query list	Hit count in genome
CID000017513	AICA	Stitch	11	126
CID000000985	Palmitate	Stitch	16	367
CID000000085	Carnitine	Stitch	14	265
CID005310993	Acipimox	Stitch	9	69
D002211	Capsaicin	CTD	11	133
CID000001424	13-HODE	Stitch	11	136

The Circadian System Is Linked to a Master Regulator of Energy Metabolism: The Transcriptional Coactivator, Peroxisome Proliferative Activated Receptor Gamma Coactivator 1 Alpha (PPARGC1A)

We used another bioinformatic tool, Genemania (GeneMANIA.org [23]), to extend the previous list with functionally similar genes that the program identifies using available genomics and proteomics data reporting weights that indicate the predictive value of each selected data set for the query. As depicted in Fig. 9.4, among the new loci predicted is the peroxisome proliferator-activated receptor gamma cofactor 1 alpha gene (*PPARG1A*, also known as *PGC1A*). The transcriptional coactivator *PPARG1A* coordinates the regulation of genes involved in energy metabolism by controlling transcriptional programs of mitochondrial biogenesis, adaptive thermogenesis, and fatty-acid betaoxidation. In fact, its tissue specificity pattern of expression is mainly located in the heart, skeletal muscle, liver, and kidney [24].



Functions legend
query genes

Networks legend
Co-expression
Co-localization
Genetic interactions
Pathway
Physical interactions
Predicted
Shared protein domains

Fig. 9.4 Genemania predictive network using interactions as depicted in the legend. New loci are rendered (*open circles*). The *arrow* indicated the PPARG cofactor 1 alpha (PPARGC1A, also known as PGC1A)

Interestingly, the protein encoded by this gene is involved in controlling blood pressure, regulating cellular cholesterol homoeostasis, and development of obesity, and altered signaling of *PPARGC1A* contributes to glucose intolerance, IR, and T2D [25, 26]. It is worthy to mention that we observed that methylation levels of the *PPARGC1A* promoter in the liver tissue of NAFLD patients are correlated with HOMA-IR and plasma fasting insulin levels [27]. Interestingly, we also observed that liver abundance of *PPARGC1A* mRNA is inversely correlated with the methylation levels of *PPARGC1A* promoter CpGs, and also with the status of peripheral IR, suggesting that methylation of, at least, three explored CpG sites in the gene promoter efficiently repressed its transcriptional activity [27]. Finally, we were able to show that mitochondrial biogenesis is reduced in the liver of NAFLD patients and is associated with peripheral IR and *PPARGC1A* promoter methylation status [27].

Expert Opinion: The Problem of the Missing Heritability

The effect of isolated genes on the disease susceptibility is in general moderate, being odds ratios (ORs) smaller than 2, at best. Even considering the sum of the effect of several risk variants combined (i.e., for seven hypothetical variants with a minor allele frequency of 0.5 and a OR of 1.5, the overall composite effect is smaller than 10 and the probability of finding individuals of such haplotypes would be less than 5% in the total population) the total variance of the phenotype explained is commonly less than 5%. A clear example can be found in the study by Li et al. [28] who showed that individuals who carried >16 risk alleles for obesity had higher BMI than those who carried <7 risk alleles but by only 1.53 BMI units and all SNPs add only 3% to the predictive value of obesity in addition to age and sex.

Then, we should considered the possibility that risk variants are not acting independently but rather by a synergistic effect, a phenomenon known as epistasis. Owing to the difficulties of the study design, few authors have reported the role of epistasis on the risk of MetS-associated phenotypes, for instance, hypertension [29], myocardial infarction and coronary artery disease [30], cholesterol levels [31], triglyceridemia [32]. In this vein, we have shown that *CLOCK* and serotonin transporter (*SLC6A4*) variants interacting with environmental factors, such as the rotating shift work have an strong effect on the overall MetS development and/or its isolated components, such as blood pressure or plasma triglycerides [33].

Recent advances in epigenomic approaches have placed the epigenetic gene regulation as a key factor in the pathogenesis of many complex disorders, mostly cancer but also other complex diseases, such as MetS, as mentioned above. In particular, epigenetic modifications can explain the mechanisms involved in the gene–environment interaction, the sexual dimorphism observed in some phenotypes, and the role of developmental programming. The most attractive aspect of the hypothesis of the impact of epigenetic changes in the etiology of the MetS-associated phenotypes is given by the nature of the epigenetic regulation, which is dynamic and subjected to either external or internal influences. Hence, while epigenetic

marks can be propagated during cell division resulting in a permanent modification of the phenotype, they are also plausible of therapeutic modifications.

The most studied mechanism of epigenetic modifications influencing a MetS-related trait is DNA methylation. For instance, DNA methylation of the *PPARGC1A* promoter in pancreatic islets from type 2 diabetic patients was associated with alterations in the insulin secretion [34]. In this vein, there is increasing evidence that prenatal environment can modify the epigenetic regulation of specific genes. We recently reported a positive correlation between maternal BMI and *PPARGC1A* promoter methylation in umbilical cord of their offspring's, suggesting a potential role of promoter *PPARGC1A* methylation in the metabolic programming of the fetus [35]. In addition, we explored whether DNA methylation of the mitochondrial transcription factor A (*TFAM*) promoter is associated with insulin resistance in adolescents with features of MS and observed a potential role of epigenetic modifications in this transcription factor in association with HOMA-IR and fasting plasma insulin levels [36]. Interestingly, because *PPARGC1A* and *TFAM* are, both, regulators of mitochondrial biogenesis, DNA methylation at their promoters may be one of cause of the mitochondrial DNA decrease we have observed in small and large for gestational age newborns as well as adolescents with insulin resistance [37, 38].

A Practical Information: How to Apply This New Concepts in the Clinical Practice

The systems biology approach is designed to analyze and integrate genomic, transcriptomic, and/or proteomic data to infer from genetic signals related pathways of disease. MetS is a complex constellation of diseases, thus it is possible to infer that common physiologic processes and molecular networks influence the risk of each intermediate phenotype. Accumulating evidence indicates that circadian desynchronization and/or alterations in circadian clock gene function have both a strong impact on the maintaining of metabolic homeostasis and cardiovascular function.

The components of the circadian system can be regarded as a network of complex interactions working either synergistically or in an integrated system that modulate the susceptibility of MetS associated phenotypes.

The most paradigmatic example in clinical practice of disruption of circadian rhythmicity is the observed in workers under rotating shift work schedules, which represents an important risk factor for the development of MetS [20, 21]. This point is particularly relevant because in industrialized nations and modern societies, as many as 20% of workers are rotating shift workers [39].

Although the epidemiological and clinical studies have provided a well-documented association among shift work schedule and MetS and cardiovascular disease, the molecular mechanisms underlying these phenomena are not yet well understood. Systems biology may shed light on novel putative pathobiological pathways associated with the disease and may also suggest novel therapeutic approaches. For instance, the interplay of serotonin, light, and the clock circadian system and

their impact on metabolic functions may provide Health professionals with new insight into the etiology and the treatment of MetS-related phenotypes in rotating shift workers.

Summary Points

- The major function of the circadian system is the internal cycling of physiologic and metabolic events to adapt cell metabolism to external cues such as sleep–awake rhythm, light–night cycle, and food availability.
- Altering the circadian rhythmicity results in pathophysiological changes inducing MetS and fat accumulation.
- Polymorphisms in the *CLOCK* and other internal clock genes (and related haplotypes) are critically involved in the genetic susceptibility to obesity
- Systems biology analysis shows that the circadian system-related genes and the predicted related proteins are highly involved in the regulation of the transcriptional activity of RNA polymerase II probably as nuclear factors "per se" or through their activity on epigenetic marks, i.e., histone deacetylases, SIRTs, etc.
- Serotonin transporter is involved in the modulation of the obesity phenotype and shows a significant interaction with the master gene *CLOCK*. These findings may have therapeutic implications.
- The circadian system is linked to several master regulator of energy metabolism. One of the most important being the transcriptional coactivator, peroxisome proliferative activated receptor gamma coactivator 1 alpha, also known as PGC1a (*PPARGC1A).* Epigenetic regulation of its activity may be important in the disease process. *PPARGC,* and other nuclear factors, are critically involved in mitochondrial function which seems to be compromised in the MetS.

Acknowledgments This study was partially supported by grants PICT 2008-1521 and 2010-0441 (Agencia Nacional de Promoción Científica y Tecnológica), and UBACYT CM04 (Universidad de Buenos Aires). S.S. and C.J.P. belong to CONICET.

References

1. Reaven GM (1998) Insulin resistance may or may not play a role in blood pressure regulation. J Intern Med 244(4):359–360
2. Sookoian S, Gianotti TF, Schuman M, Pirola CJ (2009) Gene prioritization based on biological plausibility over genome wide association studies renders new loci associated with type 2 diabetes. Genet Med 11(5):338–343
3. Sookoian S, Pirola CJ (2011) Metabolic syndrome: from the genetics to the pathophysiology. Curr Hypertens Rep 13(2):149–157
4. Turek FW, Joshu C, Kohsaka A, Lin E, Ivanova G, McDearmon E et al (2005) Obesity and metabolic syndrome in circadian Clock mutant mice. Science 308(5724):1043–1045

5. Eckel-Mahan KL, Patel VR, Mohney RP, Vignola KS, Baldi P, Sassone-Corsi P (2012) Coordination of the transcriptome and metabolome by the circadian clock. Proc Natl Acad Sci U S A 109(14):5541–5546
6. Sookoian S, Gemma C, Gianotti TF, Burgueno A, Castano G, Pirola CJ (2008) Genetic variants of Clock transcription factor are associated with individual susceptibility to obesity. Am J Clin Nutr 87(6):1606–1615
7. Scott EM, Carter AM, Grant PJ (2008) Association between polymorphisms in the Clock gene, obesity and the metabolic syndrome in man. Int J Obes (Lond) 32(4):658–662
8. Garaulet M, Corbalan MD, Madrid JA, Morales E, Baraza JC, Lee YC et al (2010) CLOCK gene is implicated in weight reduction in obese patients participating in a dietary programme based on the Mediterranean diet. Int J Obes (Lond) 34(3):516–523
9. Garaulet M, Lee YC, Shen J, Parnell LD, Arnett DK, Tsai MY et al (2009) CLOCK genetic variation and metabolic syndrome risk: modulation by monounsaturated fatty acids. Am J Clin Nutr 90(6):1466–1475
10. Sookoian S, Castano G, Gemma C, Gianotti TF, Pirola CJ (2007) Common genetic variations in CLOCK transcription factor are associated with nonalcoholic fatty liver disease. World J Gastroenterol 13(31):4242–4248
11. Maury E, Ramsey KM, Bass J (2010) Circadian rhythms and metabolic syndrome: from experimental genetics to human disease. Circ Res 106(3):447–462
12. Stubblefield JJ, Terrien J, Green CB (2012) Nocturnin: at the crossroads of clocks and metabolism. Trends Endocrinol Metab 23(7):326–333 [Epub ahead of print]
13. Barbosa-Silva A, Fontaine JF, Donnard ER, Stussi F, Ortega JM, Andrade-Navarro MA (2011) PESCADOR, a web-based tool to assist text-mining of biointeractions extracted from PubMed queries. BMC Bioinformatics 12:435–444
14. Vidal M (2005) Interactome modeling. FEBS Lett 579(8):1834–1838
15. Chen J, Bardes EE, Aronow BJ, Jegga AG (2009) ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res 37(Web Server issue): W305–W311
16. Sookoian S, Gemma C, Pirola CJ (2010) 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 99(1):80–89
17. Sookoian S, Gianotti TF, Gemma C, Burgueno AL, Pirola CJ (2010) Role of genetic variation in insulin-like growth factor 1 receptor on insulin resistance and arterial hypertension. J Hypertens 28(6):1194–1202
18. Sookoian S, Gemma C, Garcia SI, Gianotti TF, Dieuzeide G, Roussos A et al (2007) Short allele of serotonin transporter gene promoter is a risk factor for obesity in adolescents SOOKOIAN2007A. Obesity (Silver Spring) 15(2):271–276
19. Sookoian S, Gianotti TF, Gemma C, Burgueno A, Pirola CJ (2008) Contribution of the functional 5-HTTLPR variant of the slc6a4 gene to obesity risk in male adults SOOKOIAN2008. Obesity (Silver Spring) 16(2):488–491
20. Burgueno A, Gemma C, Gianotti TF, Sookoian S, Pirola CJ (2010) Increased levels of resistin in rotating shift workers: a potential mediator of cardiovascular risk associated with circadian misalignment. Atherosclerosis 210(2):625–629
21. Sookoian S, Gemma C, Fernandez GT, Burgueno A, Alvarez A, Gonzalez CD et al (2007) Effects of rotating shift work on biomarkers of metabolic syndrome and inflammation SOOKOIAN2007. J Intern Med 261(3):285–292
22. Sookoian S, Pirola CJ (2009) Shift work and subclinical atherosclerosis: recommendations for fatty liver disease detection. Atherosclerosis 207(2):346–347
23. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P et al (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 38(Web Server issue):W214–W220
24. Esterbauer H, Oberkofler H, Krempler F, Patsch W (1999) Human peroxisome proliferator activated receptor gamma coactivator 1 (PPARGC1) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression. Genomics 62(1):98–102

25. Finck BN, Kelly DP (2006) PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. J Clin Invest 116(3):615–622
26. Sookoian S, Garcia SI, Porto PI, Dieuzeide G, Gonzalez CD, Pirola CJ (2005) Peroxisome proliferator-activated receptor gamma and its coactivator-1 alpha may be associated with features of the metabolic syndrome in adolescents. J Mol Endocrinol 35(2):373–380
27. Sookoian S, Rosselli MS, Gemma C, Burgueno AL, Fernandez GT, Castano GO et al (2010) 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 52(6):1992–2000
28. Li S, Zhao JH, Luan J, Luben RN, Rodwell SA, Khaw KT et al (2010) Cumulative effects and predictive value of common obesity-susceptibility variants identified by genome-wide association studies. Am J Clin Nutr 91(1):184–190
29. Tomaszewski M, Charchar FJ, Lacka B, Pesonen U, Wang WY, Zukowska-Szczechowska E et al (2004) Epistatic interaction between beta2-adrenergic receptor and neuropeptide Y genes influences LDL-cholesterol in hypertension. Hypertension 44(5):689–694
30. Baessler A, Fischer M, Mayer B, Koehler M, Wiedmann S, Stark K et al (2007) Epistatic interaction between haplotypes of the ghrelin ligand and receptor genes influence susceptibility to myocardial infarction and coronary artery disease. Hum Mol Genet 16(8):887–899
31. Isaacs A, Aulchenko YS, Hofman A, Sijbrands EJ, Sayed-Tabatabaei FA, Klungel OH et al (2007) Epistatic effect of cholesteryl ester transfer protein and hepatic lipase on serum high-density lipoprotein cholesterol levels. J Clin Endocrinol Metab 92(7):2680–2687
32. Brisson D, St-Pierre J, Santure M, Hudson TJ, Despres JP, Vohl MC et al (2007) Genetic epistasis in the VLDL catabolic pathway is associated with deleterious variations on triglyceridemia in obese subjects. Int J Obes (Lond) 31(8):1325–1333
33. Sookoian S, Gianotti TF, Burgueno A, Pirola CJ (2010) Gene–gene interaction between serotonin transporter (SLC6A4) and clock modulates the risk of metabolic syndrome in rotating shiftworkers. Chronobiol Int 27(6):1202–1218
34. Ling C, Del GS, Lupi R, Ronn T, Granhall C, Luthman H et al (2008) Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. Diabetologia 51(4):615–622
35. Gemma C, Sookoian S, Alvarinas J, Garcia SI, Quintana L, Kanevsky D et al (2009) Maternal pregestational BMI is associated with methylation of the PPARGC1A promoter in newborns. Obesity (Silver Spring) 17(5):1032–1039
36. Gemma C, Sookoian S, Dieuzeide G, Garcia SI, Gianotti TF, Gonzalez CD et al (2010) Methylation of TFAM gene promoter in peripheral white blood cells is associated with insulin resistance in adolescents. Mol Genet Metab 100(1):83–87
37. Gemma C, Sookoian S, Alvarinas J, Garcia SI, Quintana L, Kanevsky D et al (2006) Mitochondrial DNA Depletion in Small- and Large-for-Gestational-Age Newborns. Obesity (Silver Spring) 14(12):2193–2199
38. Gianotti TF, Sookoian S, Dieuzeide G, Garcia SI, Gemma C, Gonzalez CD et al (2008) A Decreased Mitochondrial DNA Content Is Related to Insulin Resistance in Adolescents. Obesity 16(7):1591–1595

List of Bioinformatic Resources

PESCADOR: available at: http://cbdm.mdc-berlin.de/tools/pescador/
ToppGene Suite: available at: http://toppgene.cchmc.org
GeneMANIA: available at: http://www.genemania.org
KEGG: Kyoto Encyclopedia of Genes and Genomes: available at: http://www.genome.jp/kegg