Polymorphisms in the *SRNPN* gene are associated with obesity susceptibility among Spanish population

David Albuquerque[1,2], Licínio Manco[1], Luz M González[2], Guillermo Gervasini[3], Goitzane Marcaida Benito[2,4], Juan R González[5,6,7*], Raquel Rodríguez-López[2,4*]

[1]Research Center for Anthropology and Health (CIAS), University of Coimbra, Coimbra, Portugal

[2]Genomics group, Fundación Investigación Hospital General Universitario de Valencia, Valencia, Spain

[3]Department of Medical & Surgical Therapeutics, Division of Pharmacology, Medical School, University of Extremadura, Badajoz, Spain.

[4]Laboratory of Molecular Genetics, Clinical Analysis Service, Hospital Universitario General de Valencia, Valencia, Spain

[5]ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain.

[6]CIBER Epidemiología y Salud Pública (CIBERESP).

[7]Universitat Pompeu Fabra (UPF), Barcelona, Spain

*These authors contributed equally to this work

Running title:

Association of *SNRPN* gene with common obesity

Corresponding authors:
Raquel Rodríguez-López, MD, PhD
Laboratory of Molecular Genetics, Clinical Analysis Service
General University Hospital, Valencia, Spain
E-mail address: rodriguez_raqlop@gva.es
Tel.: 0034 963 131 800 – 437 317
Fax: 0034 963 131 979

Juan R. González, PhD
Bioinformatic Research Group in Epidemiology (BRGE)
Barcelona Institute for Global Health
Avd Dr Aiguader, 88 | Barcelona 08003
e-mail address: juanr.gonzalez@isglobal.org
Telf: +34 932 147 327

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jgm.2956

ABSTRACT

Background: *SNRPN*, which codes for the RNA-binding SmN protein, is a candidate gene for Prader-Willi syndrome. One characteristic of this neuroendocrine disorder is hyperphagia resulting in extreme obesity later in life. In this study we aim to assess whether variability within this gene could be implicated in obesity susceptibility.

Material and methods: A case-control study was performed including 265 unrelated patients with non-syndromic and early-onset severe obesity, belonging to high risk obesity families from Spanish ancestry; 184 healthy control individuals were included representative of the same genetic background and sex-matched. Forty-nine single nucleotide polymorphisms (SNPs) spanning the entire SNRPN gene were selected and genotyped by using Sequenom MassARRAY platform.

Results: The four SNPs rs12905653, rs752874, rs1391516, and rs2047433 were found nominally associated with obesity ($p<0.03$). The diversity haplotype distribution among cases and controls identified the combination rs12905653-T/rs8028366-A/rs4028395-T strongly and inversely associated with obesity (OR=0.49; p=0.0006). A genetic risk score was built based on rs12905653, rs1391516, and rs2047433 SNPs and each unit higher GRS increase the obesity risk by 49% (OR=1.49; CI95%: 1.24-1.80).

Conclusions: To our knowledge, this is the first study reporting an association between variability in the *SNRPN* gene with the risk of being obese. Interestingly, it was the major allele of each SNP, which was found associated with the risk of weight gain. Further studies analyzing this locus, and the possible additive deleterious capability of SNPs combinations, could be useful to demonstrate the obesity development.

Key words: obesity, genetic susceptibility, *SNRPN* gene, Spanish population, BMI, case-control study.

INTRODUCTION

The worldwide increased prevalence of obesity and related co-morbidities has reached epidemic proportions, becoming one of the major public health problems [1]. While genetic factors undoubtedly contribute to individual obesity susceptibility, the obesity-associated single nucleotide polymorphisms (SNPs) that have been identified to date explain less than 3% of the inherited susceptibility to develop an obese phenotype [2]. This situation suggests that additional genetic *loci* predisposing to weight gain remain to be discovered.

Genome-wide association studies (GWAS) have been successful in identifying genetic variants associated with complex traits. In the last decade, GWAS have identified more than 50 *loci* associated with common obesity [2]. However, the candidate gene approach is still used in association analyses, as it is more specific to identify potential causative gene variants for a particular trait [3]. We hypothesize that investigating syndromic diseases whereof extreme obesity is part of their phenotype may also help in identifying new variants for weight gain.

The Prader-Willi syndrome (PWS) is considered one of the most common genetic syndromes causing morbid obesity in children [4]. This condition is caused by imprinted genes within the 15q11-q13 region expressed exclusively from the paternal chromosome [4, 5]. The small nuclear ribonucleoprotein polypeptide N (*SNRPN*) gene, encoding the RNA-binding SmN splicing factor involved in RNA processing, is one of the best described genes linked to PWS manifestation [6]. The *SNRPN* gene (15q11.2) includes ten exons that span about 155 kb [7]. It is expressed in a tissue-specific manner, with the higher expression levels in the adult brain and heart [8].

In this study, we aimed to assess whether common variants in the *SNRPN* gene are associated with obesity. This hypothesis was considered attending on several facts: i) SNPs

located in the *SNRPN* gene have not been studied in relation to their association with common obesity; ii) *SNRPN* gene maps within the imprinted gene cluster (15q11-q13), which is associated with clinical manifestations of PWS; iii) the *SNRPN* gene is considered a putative candidate for obesity susceptibility since Chen *et al*. [9] identified three copy number variations (CNVs) overlapping this gene which were significantly associated with body fat mass.

Patients and methods

Patients and controls

A total of 265 unrelated, non-syndromic and severely obese Spanish individuals, 44.6% were females, were recruited from endocrine and metabolic health service at the Infanta Cristina Hospital (Badajoz, Spain). All of the enrolled patients reached grade 3 overweight with early-onset severe obesity at a medium age of 6 years, weight greater than (mean +3 SD) before 14 years of age, and who referred at least two other morbid obesity cases among first- or second- degree relatives. We consider the existence of grade 3 overweight to be a BMI greater than or equal to 40 kg/m^2; in children this corresponds to a BMI greater than the 95^{th} percentile for age-matched and sex-matched control subjects from our population data. To reduce required sample size, all cases were selected and well-characterized from high risk obesity families screened as a part of a previous study to detect a genetic risk profile of obesity [10]. The control group comprises 184 healthy Spanish adults, 54.1% were females. The controls were recruited attending at primary healthcare services for routine clinical analysis in Badajoz (Spain). These were asked through nurse if they wanted to participate in the study. They were randomly chosen only to matched age and gender as the cases. All individuals enrolled were from Extremadura (Western Spanish region).

The study protocol was approved by the Ethical Committee of the Infanta Cristina Hospital, and was conducted in accordance with the Declaration of Helsinki and its subsequent revisions. Written informed consent was previously obtained from all patients prior to enrolment.

Tag SNPs selection and genotyping

For this study, we analyzed a 155-kb region of chromosome 15 (spanning 24,823,637 to 24,978,723 of human reference sequence GRCh38:CM000677.2, encompassing the candidate *SNRPN* gene, potentially involved in genetic susceptibility risk to severe obesity. To minimize the burden of genotyping multiple variants localized in the candidate *SNRPN* gene, we used a tagging SNP approach to capture all gene-specific regions for analysis. Forty-nine tag SNPs from a total of 227 with minor allele frequencies (MAF) >10%, and pair-wise tagging less than r^2>0.80 (HapMap project data) were selected. None of these 49 common SNPs were predicted to generate amino acid substitutions and all of them were located in intragenic regions.

This candidate genomic region displays high linkage disequilibrium (LD), which is characterized by strong association between alleles, low haplotype diversity and low recombination rates. Data regarding common SNPs and LD values in Caucasians (CEU) were obtained from Ensembl (http://www.ensembl.org) and Hapmap (http://www.hapmap.org) databases. Tag SNPs were assigned using the Haploview software package (http://www.broad.mit.edu/mpg/haploview/). Ancestral alleles were inferred from the Ensembl EPO 8 primate whole genome alignments (http://www.ensembl.org).

Genomic DNAs were extracted from peripheral blood samples of the study subjects using the QIAamp DNA Blood Kit (Qiagen, GmbH, Hilden, Germany) following the manufacturer's instructions. PCRs and genotyping analyses were performed using primer

extension chemistry of the iPLEX® assay and the mass spectrometry-based platform of Mass ARRAY MALDI-TOF (Sequenom Inc., San Diego, USA), according to the manufacturer's instructions for the multiplex reactions. The genotype calls were generated and manually checked by two individuals separately using Sequenom TYPER software v.4 (Sequenom Inc., San Diego, USA). Duplicate samples and negative controls were included across the plates to ensure accuracy of genotyping. In addition, to confirm the genotyping quality, 10% of random samples were selected and re-genotyped for all SNPs on the Sequenom platform according to manufacturer's instructions, and revealed 100% concordance. Duplicate samples across plates similarly showed no discrepant calls. Study design and conditions for the iPLEX® assay are available upon request.

Statistical analyses

The allelic and genotypic frequencies of all SNPs were estimated by direct counting. Genotype distributions were tested at each biallelic SNP marker for departure from Hardy-Weinberg equilibrium (HWE) by using an exact test [11]. Logistic regression under an additive mode of inheritance, allowing for analysis of individual SNPs, was performed using the likelihood ratio test (LRT), adjusted for age and gender. All these statistical analyses were carried out with the *SNPassoc* R package [12]. Haplotypes associations with obesity were performed with the *haplo.stats* R package [13]. In order to refine the haplotype construction and to identify a potential causative core region, we used an unbiased sliding window of SNPs approach to determine those blocks associated with obesity. For each window, haplotype frequencies in cases and controls were estimated using an EM algorithm and the permutation procedure was used to estimate the significance of the best result (based on 1000 permutations). LD plots were created by computing the r^2 statistic. Power to test the association of haplotypes with obesity status was performed using *haplo.stats* R library that

implements the methods described in [14]. In addition, a genetic risk score (GRS) was computed by adding the number of risk alleles (0, 1 or 2) across selected SNPs and its predictive value for obesity risk was calculated as the area under the curve (AUC) [15]. A logistic regression model was also used to estimate the risk per each unit allele increase.

eQTL analysis

We have performed an *in silico* data analysis to investigate whether our GRS has any transcriptomic effect. To address this question, gene expression data of RNA from lymphoblastoid cell lines of CEU individuals generated through the Project E-MTAB-198 were obtained from the European Bioinformatics Institute at EMBL (http://www.ebi.ac.uk/arrayexpress/). Only data from 105 unrelated individuals were analyzed, removing samples from family relatives. The eQTL analysis was performed by creating the GRS using SNP data that were obtained from HapMap project. Linear regression models were used to assess association between the GRS and gene expression of those HapMap samples using limma Bioconductor package [16]. By using these models, we are able to test whether having more risk alleles of our proposed GRS increase or decrease gene expression.

RESULTS

We performed a case-control study to identify the association of SNPs with obesity within the *SNRPN* gene in a sample of Spanish patients. Descriptive characteristics of the study sample have been previously reported [10].

A discovery analysis was carried out by studying 49 tag SNPs mapping the entire *SRNPR* gene region. The genotyping success rate for each SNP varied between 96.0% and 99.3%. Minor allele frequencies and departure from HWE for all SNPs in the control group are

shown in Supplementary Table 1. These frequencies were in accordance with those found in the HapMap IBS population (http://www.ensembl.org/). Genotype frequencies among the control group were in agreement with HWE (p>0.05), except for two SNPs (rs11161160 and rs2647364), which were consequently excluded from further analyses.

Univariate analyses under an additive genetic model revealed that four SNPs (rs12905653, rs752874, rs1391516 and rs2047433) showed a protective effect of the minor allele against the development of obesity, with p values ranging from 0.0012 to 0.0295 (Figure 1, Table 1). Supplementary Table 2 shows all univariate analyses for the overall SNPs analysed.

Our study has a 83% of power to find an odd ratio of 0.49 with frequency of haplotypes described in Table 2, assuming a type-I error of 5% and a prevalence of 20% of obesity [17]. The haplotype analysis by using the sliding window approach showed that the best marker combination contained three adjacent markers: rs12905653, rs8028366, and rs4028395. Of note, rs12905653 had also been the SNP with the strongest association with obesity observed in the previous univariate analyses. These three SNPs are located between introns 1 and 2 of the *SNRPN* gene spanning chr15 at 25074767-25092266 bp (Figure 1). A total of four common haplotypes were statistically derived with frequencies ≥10%. Testing for all haplotype effects, a significant association for this 3-SNP sliding window was found towards protection against obesity (χ^2_4=18.04; p=0.0012; Table 2). In particular, the TAT haplotype (with a frequency of 15%) showed a highly significant protective effect (OR=0.49, 95% CI: 0.32-0.73, p=0.0006) compared with the reference CCA haplotype (frequency of 38%).

In order to create the GRS, we used a multivariate model with stepwise procedure based on AUC which found three SNPs being selected more often than any others in the simulation to predict the probability of being obese. This risk increased by 49% for each unit in the number of risk alleles (OR=1.49; 95% CI: 1.24-1.80). The three *loci* GRS for obesity

conferred a maximum risk for homozygous individuals rs12905653 (CC), rs1391516 (TT), and rs2047433 (GG).

Finally, we explored whether the alleles contained in the GRS might have gene regulatory effects. Fort this, we analyzed our GRS for gene expression in whole blood samples based on 105 CEU HapMap individuals in the region of interest. We observed one significant eQTL regulating *MKRN3* expression (p=0.0185) on chr15:23,813,273-23,813,322, only 1 Mb chromosomal position, centromeric from *SNRPN* gene. This implies that having more risk alleles increases expression of *MKRN3* gene (Supplementary Table 3). However, after FDR correction of the results, the significance of association decreased to p=0.0545.

DISCUSSION

In the last decade, an increasing number of GWA studies in individuals affected by common obesity, have led to the identification of hundreds of *loci* affecting the risk to weight gain [2]. However, we should not forget that *a priori* hypotheses are needed to establish putative candidate genes. Highly penetrant genomic *de novo* disorders in affected individuals with intellectual disability and/or congenital malformations, have allowed to implicate causative genes in other common clinical features. Combination of higher SNPs/CNVs microarrays, with genotyping of larger disease cohorts and controls, has an interesting potential to evidence the risk genetic bases of traits as overweight and obesity. In 2011, Chen *et al.* [9] reported three CNVs at the PWS region (15q11.2-q13) that were associated with body fat mass in the general Caucasian population. These CNVs are located near the necdin homolog (*NDN*) and chromosome 15 open reading frame 2 (*C15orf2*) genes, with higher copy numbers resulting in an increase of 5.08-9.77 kg in body fat mass. Other genes have been identified with a higher impact on PWS, whereof morbid obesity is one of the distinctive features. This chromosomal region is characterized by an imprinting control area

that regulates parent-of-origin of *SNRPN*, *NDN*, *MAGEL2*, and *MKRN3* genes, which are expressed exclusively from the paternal allele. Two other genes, *UBE3A* and *ATP10A*, are expressed exclusively from the maternal allele in specific tissues [18, 19]. The imprinting center region maps in part to the promoter and first exon of *SNRPN* gene [20].

We identified four SNPs, namely rs12905653, rs752874, rs1391516 and rs2047433, located within the first and second introns of the *SNRPN* gene, which were nominally associated (p values ranged from 0.0012 to 0.0295) with increased risk to develop an obese phenotype. After performing haplotype analysis we identified the rs12905653/rs8028366/rs4028395 TAT haplotype in a high LD block to be robustly associated with decreased risk of obesity (p=0.0006). Indeed, it was the rs12905653 major C-allele which showed higher frequency among the cases (87.8%) than among control group (79.4%). This result suggests that the rs12905653 minor T-allele has a protective effect of obesity against C-allele (p=0.0012). Only the SNP rs12905653 included in this haplotype was previously found nominally associated with the susceptibility risk to develop obesity, confirming the potential role of the rs12905653 major C-allele on risk of obesity in our sample. Interestingly, out of the four SNPs found nominally associated with obesity, the minor alleles rs12905653 (T) and rs752874 (T) associated with protection against obesity are actually the ancestor alleles (based on Ensembl's 8 primates alignment). Regarding SNPs rs1391516 and rs2047433, the major alleles T and G, respectively, are the ancestor alleles (based on Ensembl's 8 primates alignment) and were associated with obesity risk. Due to the lack of data concerning *SNRPN* SNPs analysis in the literature we cannot infer about allelic distributions. However, when comparing our allelic frequencies with those present in HapMap databases, these are similar with those reported for the Iberian population (IBS, Ensembl).

Polygenic obesity apart environmental factors is most likely due to a cumulative set of several SNPs located within different genes, each of them contributing with a specific signal [21]. Moreover, the creation of a genetic risk score based on these SNPs may constitute a good predictor for the risk to develop an obese phenotype [22]. We performed a GRS analysis based on the 47 *SNRPN* SNPs which allowed a combination of three SNPs (rs12905653, rs1391516, and rs2047433) to better predict the probability of being obese. The SNP rs12905653 was again present in this analysis confirming its potential relation with obesity. The risk for obesity increased according to the number of risk alleles (49% per each additional risk allele).

In a previous study using the same series of individuals, six SNPs located in or near the *FTO*, *TFAP2B*, *SEC16B*, *ETV5* and *SH2B1*genes have been found associated with obesity [22], confirming previous findings regarding these *loci*. Thus, the sample size and constitution of our study sample have sufficient statistical power detection in case-control studies. Furthermore, seemed curious that no patient carrying this new predictive score of obesity, carried the previous six *loci* GRS for obesity, which was identified in this series [10].

Our results suggest a contribution of the *SNRPN* SNPs to a predisposition to obesity. The *SNRPN* gene is located within an imprinted gene cluster, exclusively expressed from the paternal allele [4, 5]. The absence of the paternally expressed gene may not have an identifiable coding capacity. However, one limitation of the study is that we cannot have the possibility to perform a DNA methylation pattern analysis in our series. Thus, we did not know if the gene was or not imprinted in both or one parent's alleles. It could be interesting further analysis one this way. In this regard, a previous study performed by Carobin *et al*. [23] to investigate the possible differences in *SNRPN* gene methylation profiles in non-obese and obese individuals did not find any significant results. However, Soubry *et al*. [24] suggest that a contribution of paternal obesity can change epigenetic programming in a small fraction

of sperm cells. Their results indicate that male overweight/obese status could be traceable in the sperm epigenome.

The combination between SNPs and eQTLs may be important to confirm the relation of the SNPs with a trait [25]. Indeed, if one allele is more frequently present in cases than in controls and at the same time it impacts the expression of a nearby gene, this may probably establish causality for this variant. We observed a significant association between our GRS and the *MKRN3* gene expression located 1.25 Mb downstream the *SNRPN* gene. However, after several corrections the *p*-value decreased to a borderline significance.

To the best of our knowledge, this study is the first one to investigate the association between common variants and derived haplotypes of the selected *SNRPN* candidate gene in a sample of severely obese individuals. Our series of cases is composed by individuals from families with a high proportion of obese individuals (≥ 3 relatives). This, together with the inclusion of extreme phenotypes, such as a mean age of 6 years for severe obesity onset, increases the power to detect variants associated with obesity in our otherwise moderate sample size. We also consider that the use of uniform criteria in the series is more important and efficient to ensure a comprehensive assessment of personal and familial phenotype of each proband. Furthermore, the analysis of haplotypes with the grouping and interaction of several variants has claimed to be superior to individual SNP analysis [18].

In conclusion, our study expanded the current knowledge on the genetic basis of obesity by identifying new SNPs associated with obesity risk in a sample of Spanish individuals with severe obesity. Further studies analyzing this locus are needed to improve our knowledge of the relative contribution of PWS-related genes in human obesity. Additional studies based on the interplay between genetic and epigenetic analyses on the 15q11-q13 region are required to unravel new insights on the obesity condition.

Acknowledgments

The authors are very grateful to the patients who participated in our study.

Conflict of Interests

All the authors recognize and disclose to have no conflict of interest to declare.

Funding

This work was supported by the ISCIII, Madrid for Financial support (PI070594) and Fundesalud, Junta de Extremadura (08PRI008). David Albuquerque is recipient of a post-doctoral grant (SFRH/BPD/109043/2015) from *Fundação para a Ciência e a Tecnologia* (FCT, Portugal).

REFERENCES

1. Singh G, Danaei G, Farzadfar F, et al (2016) Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. Lancet 387:1377–1396.
2. Albuquerque D, Stice E, Rodríguez-López R, et al (2015) Current review of genetics of human obesity: from molecular mechanisms to an evolutionary perspective. Mol Genet Genomics 290:1191–221.
3. Tabor HK, Risch NJ, Myers RM (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 3:391–397. doi: 10.1038/nrg796
4. Butler MG (2011) Prader-Willi Syndrome: Obesity due to Genomic Imprinting. Curr Genomics 12:204–15. doi: 10.2174/138920211795677877
5. Cassidy SB, Dykens E, Williams CA (2000) Prader-Willi and Angelman syndromes: sister imprinted disorders. Am J Med Genet 97:136–46.
6. Özçelik T, Leff S, Robinson W, et al (1992) Small nuclear ribonucleoprotein polypeptide N (SNRPN), an expressed gene in the Prader–Willi syndrome critical region. Nat Genet 2:265–269.
7. Fridman C, Koiffmann CP (2000) Genomic imprinting: genetic mechanisms and phenotypic consequences in Prader-Willi and Angelman syndromes. Genet Mol Biol 23:715–724.
8. McAllister G, Amara SG, Lerner MR (1988) Tissue-specific expression and cDNA cloning of small nuclear ribonucleoprotein-associated polypeptide N. Proc Natl Acad Sci USA 85:5296–300.
9. Chen Y, Liu Y-J, Pei Y-F, et al (2011) Copy number variations at the Prader-Willi syndrome region on chromosome 15 and associations with obesity in whites. Obesity (Silver Spring) 19:1229–34.

10. González JR, González-Carpio M, Hernández-Sáez R, et al (2012) FTO risk haplotype among early onset and severe obesity cases in a population of western Spain. Obesity (Silver Spring) 20:909–15.
11. Wigginton JE, Cutler DJ, Abecasis GR (2005) A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 76:887–93.
12. González JR, Armengol L, Solé X, et al (2007) SNPassoc: an R package to perform whole genome association studies. Bioinformatics 23:644–5.
13. Lake SL, Lyon H, Tantisira K, et al (2003) Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum Hered 55:56–65.
14. Schaid D (2006) Power and Sample Size for Testing Associations of Haplotypes with Complex Traits. Ann Hum Genet 70:116–130.
15. Cooke Bailey J, Igo R (2016) Genetic risk scores. Curr Protoc Hum Genet 91:1.29.1-1.29.9.
16. Ritchie ME, Phipson B, Wu D, et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47.
17. Aranceta-Bartrina J, Pérez-Rodrigo C, Alberdi-Aresti G, et al (2016) Prevalence of General Obesity and Abdominal Obesity in the Spanish Adult Population (Aged 25–64 Years) 2014–2015: The ENPE Study. Rev Española Cardiol 69:579–587.
18. Horsthemke B, Wagstaff J (2008) Mechanisms of imprinting of the Prader-Willi/Angelman region. Am J Med Genet A 146A:2041–52. doi: 10.1002/ajmg.a.32364
19. Herzing LB, Kim SJ, Cook EH, et al (2001) The human aminophospholipid-transporting ATPase gene ATP10C maps adjacent to UBE3A and exhibits similar imprinted expression. Am J Hum Genet 68:1501–5.
20. Yang T, Adamson TE, Resnick JL, et al (1998) A mouse model for Prader-Willi syndrome imprinting-centre mutations. Nat Genet 19:25–31.
21. Li S, Zhao JH, Luan J, et al (2010) Cumulative effects and predictive value of common obesity-susceptibility variants identified by genome-wide association studies. Am J Clin Nutr 91:184–90.

22. González JR, Estévez MN, Giralt PS, et al (2014) Genetic risk profiles for a childhood with severely overweight. Pediatr Obes 9:272–80.

23. Carobin N V, Rubatino FVM, Freitas ML, et al (2015) Methylation profile of SNRPN gene and its correlation with weight and chronological age. Genet Mol Res 14:13791–8.

24. Soubry A, Guo L, Huang Z, et al (2016) Obesity-related DNA methylation at imprinted genes in human sperm: Results from the TIEGER study. Clin Epigenetics 8:51.

25. Nicolae DL, Gamazon E, Zhang W, et al (2010) Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS. PLoS Genet 6:e1000888.

Table 1. Univariate analysis of analyzed SNPs with positive association within the *SNRPN* gene.

SNP ID	Position	Genotype	Controls		Cases		OR (CI 95%)	*p*-value
			n	**%**	***n***	**%**		
rs12905653	15:24829620	CC	115	63.2	193	77.5		
		CT	59	32.4	51	20.5		
		TT	8	4.4	5	2.0	0.55 (0.38-0.79)	0.0012
rs752874	15: 24856926	CC	42	23.3	79	29.9		
		CT	78	43.3	130	49.2		
		TT	60	33.4	55	20.9	0.70 (0.53-0.91)	0.0066
rs1391516	15:24861936	TT	65	36.1	114	43.7		
		CT	81	45.0	117	44.8		
		CC	34	18.9	30	11.5	0.73 (0.56-0.97)	0.0264
rs2047433	15:24883191	GG	65	36.5	126	47.9		
		GT	83	46.6	102	38.8		
		TT	30	16.9	35	13.3	0.74 (0.57-0.97)	0.0295

Abbreviations: SNP ID, single nucleotide polymorphism identification; OR, odds ratio (for the minor allele); CI, confidence interval; *n*, total number of genotyped individuals.

OR and *p*-value obtained by logistic regression under an additive model, and adjusted for age and sex.

Table 2. Haplotype frequencies, odds ratio (OR) and confidence interval at 95% (CI 95%) for the association between obesity and haplotypes created using SNPs rs12905653, rs8028366, rs4028395 that were selected using sliding window approach.

Haplotype	Frequency	OR	95% CI.	*p*-value
CAA	0.38	Reference		
CAT	0.28	0.88	0.61-1.27	0.4906
CTA	0.18	0.70	0.47-1.06	0.0929
TAT	0.15	0.49	0.32-0.73	**0.0006**
rare haplotype	0.02	5.46	0.68-43.97	0.1110

Abbreviations: OR, odd ratio; 95% CI, confidence interval; *p*-value significant ($p \leq 0.05$) in bold.

Rare haplotypes those with a frequency <5% were collapsed in a unique category.

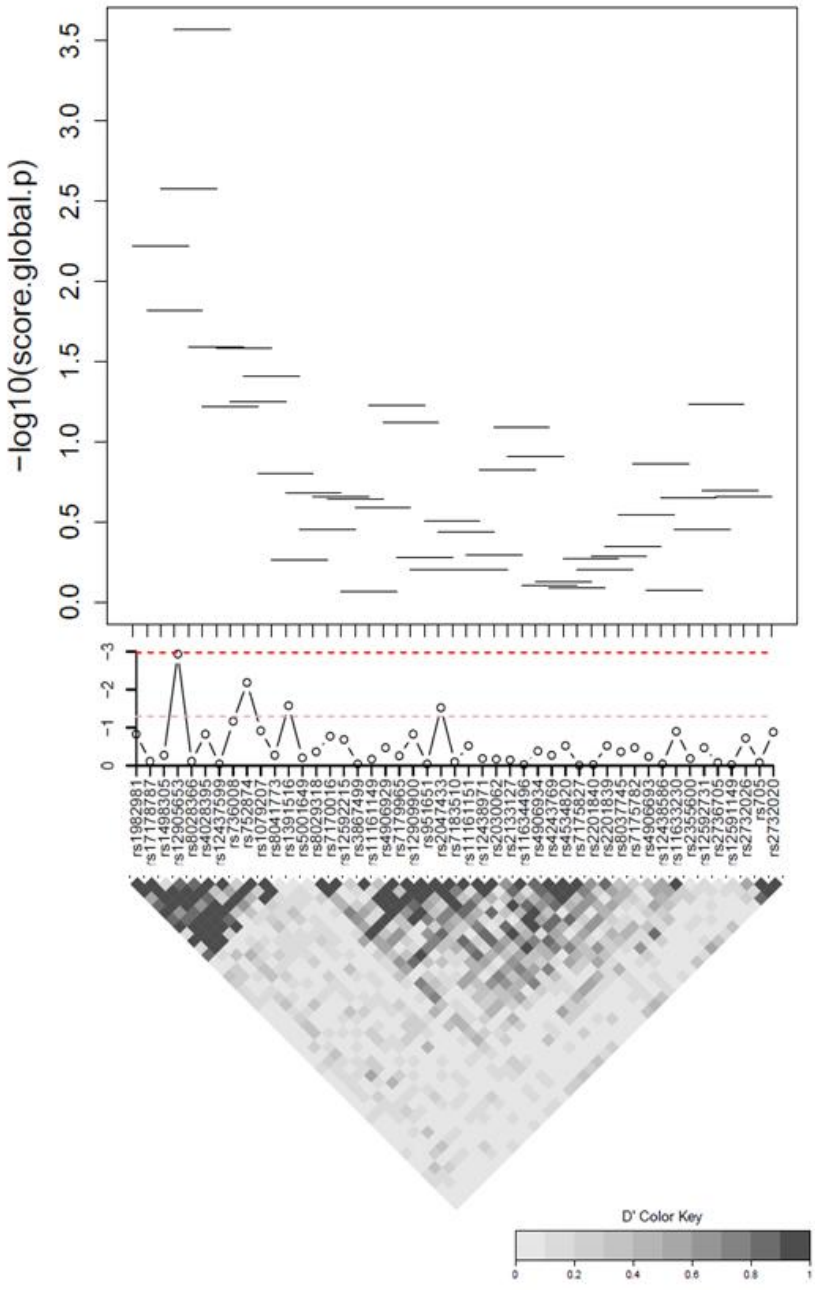

Figure 1. Bottom panel shows the linkage disequilibrium pattern among 47 selected SNPs. The middle panel gives the *p* values of association between SNPs and obesity. The highest dashed line corresponds to the threshold for the significance level after correcting for multiple comparisons. The top panel illustrates the sliding window approach.