Title: Association between *PTPN1* polymorphisms and obesity-related phenotypes in European adolescents: influence of physical activity.

Running head: *PTPN1* polymorphisms and obesity-related phenotypes

List of authors: Diego F. Salazar-Tortosa[1,2], Idoia Labayen[3], Marcela González-Gross[4], Miguel Seral-Cortes[5,6,7,8], Luis A. Moreno[5,6,7,9], Augusto G. Zapico[10], Kurt Widhalm[11], Aline Meirhaeghe[12], David Enard[2], Jonatan R. Ruiz[1,13,14]

Institutions:

[1]PROFITH 'PROmoting FITness and Health through physical activity' research group, Sport and Health University Research Institute (iMUDS), University of Granada, Granada, Spain.

[2]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona.

[3]Institute for Innovation & Sustainable Development in Food Chain (IS-FOOD), Department of Health Sciences, Navarra's Health Research Institute (IdiSNA), Public University of Navarra, Pamplona, Spain.

[4]Department of Health and Human Performance, Technical University of Madrid, Madrid, Spain.

[5]Growth, Exercise, Nutrition and Development (GENUD) Research Group, University of Zaragoza, 50009 Zaragoza, Spain.

[6]Instituto Agroalimentario de Aragón (IA2), 50013, Zaragoza, Spain.

[7]Instituto de Investigación Sanitaria Aragón (IIS Aragón), 50009 Zaragoza, Spain.

[8]Department of Psychiatry and Nursing, Faculty of Health Sciences, Universidad de Zaragoza, Spain.

[9]Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición (CIBERObn), Instituto de Salud Carlos III, 28029 Madrid, Spain.

[10]School of Education, Complutense University of Madrid, Madrid, Spain.

[11]Department of Pediatrics, Division of Clinical Nutrition, Medical University of Vienna, Währinger Gürtel 18-20, A-1090 Wien, Austria.

[12]Inserm, Institut Pasteur de Lille, University Lille, UMR1167-RID-AGE-Risk factors and molecular determinants of aging-related diseases, Lille, France.

[13]Department of Physical Education and Sport, Faculty of Sport Sciences, University of Granada, Granada, Spain.

[14]Instituto de Investigación Biosanitaria, ibs.Granada, Granada, Spain.

Corresponding author: Diego F. Salazar-Tortosa; address: Department of Ecology and Evolutionary Biology, University of Arizona, 1041 E Lowell Street 85719, Tucson, Arizona, USA; e-mail: dftortosa@arizona.edu; Tel: +1 (520) 6214871.

Contact: Jonatan R. Ruiz, e-mail: ruizj@ugr.es, Tel: +34 (958) 242754.

Category of study: Clinical study.

Impact:

- Using gene-phenotype and gene*environment analyses, we detected associations between polymorphisms of the *Protein Tyrosine Phosphatase-N1* (*PTPN1*) gene

and obesity-related phenotypes, suggesting a mechanism that can be modulated by physical activity.

- This study shows that genetic variability of *PTPN1* is associated with adiposity, while physical activity seems to modulate the genetic predisposition.

- This brings insights about the mechanisms by which physical activity positively influences obesity.

Abstract

Background: To study the associations of *Protein Tyrosine Phosphatase-N1* (*PTPN1*) polymorphisms with obesity-related phenotypes in European adolescents, and the influence of physical activity on these relationships.

Methods: Five polymorphisms of *PTPN1* were genotyped in 1,057 European adolescents (12-18 years old). We measured several phenotypes related to obesity, such as adiposity markers, and biochemical and clinical parameters. Physical activity was objectively measured by accelerometry.

Results: The T, A, T, T and G alleles of the rs6067472, rs10485614, rs2143511, rs6020608 and rs968701 polymorphisms, respectively, were associated with lower levels of obesity-related phenotypes (i.e., body mass index, body fat percentage, hip circumference, fat mass index, systolic blood pressure and leptin) in European adolescents. In addition, the TATTG haplotype was associated with lower body fat percentage and fat mass index compared to the AACCA haplotype. Finally, when physical activity levels were considered, alleles of the rs6067472, rs2143511, rs6020608 and rs968701 polymorphisms were only associated with lower adiposity in active adolescents.

Conclusions: *PTPN1* polymorphisms were associated with adiposity in European adolescents. Specifically, alleles of these polymorphisms were associated with lower adiposity only in physically active adolescents. Therefore, meeting the recommendations of daily physical activity may reduce obesity risk by modulating the genetic predisposition to obesity.

Introduction

Obesity is an important public health issue at a global scale, especially in young people[1]. This disorder possibly results from the complex interaction between genetic and lifestyle factors, being its proximate cause a chronic excess of energy intake over expenditure. Several factors can affect the balance between energy intake and expenditure. For example, basal metabolism, physical activity, thermoregulation and digestive processes influence energy expenditure, while hunger and satiety impact energy intake[2,3].

Genetic predisposition to obesity can be monogenic. This is a form of obesity characterized by its severity and early onset, being caused by a single-gene mutation with small or no modulation by environmental factors[4]. For example, this type of obesity has been described for mutations in genes involved in the control of the appetite and satiety center like *leptin* or *pro-opiomelacortin*[5]. Nevertheless, the most common form of obesity is polygenic, which is caused by the interaction between hundreds of genetic variants and environmental factors[4,6]. Physical activity is a known lifestyle factor that affects health status and can modulate genetic risk[7–9]. Regular physical activity can improve several obesity-related traits like blood pressure, dyslipidemia or sensitivity to insulin[10]. Current evidence suggests that physical activity could attenuate the genetic risk of obesity. For example, a meta-analysis of 200,452 European adults showed that physical activity may attenuate the negative effect of *FTO* polymorphisms on obesity[11]. In adolescents, we reported that achieving the daily physical activity recommendations for this age category (≥ 60 minutes/day of moderate to vigorous physical activity)[12] may reduce the genetic predisposition to obesity associated with the *FTO* rs9939609[7], *UCP1* rs2071415[8] and *LPL* rs1534649 and rs258[13] polymorphisms.

Protein tyrosine phosphatases (PTPs) are a group of enzymes implicated in the dephosphorylation of tyrosine (Tyr)-phosphorylated proteins[14]. PTPN1 was the first

protein-tyrosine-phosphatase discovered, being expressed in several tissues and involved in several signal transduction pathways. PTPN1 plays a key role in insulin signaling by dephosphorylating the insulin receptor and the insulin receptor substrate (IRS-1), which leads to downregulate its downstream signaling components[15]. PTPN1 also participates in the downregulation of leptin signaling by dephosphorylating molecules upstream in this pathway[16,17]. In addition, this protein has been related to adiposity, as evidence coming from knockout mice suggests that *Ptpn1* deletion provides resistance to weight gain[15,18]. Several genetic association studies have also associated *PTPN1* polymorphisms with obesity-related phenotypes like insulin sensitivity or glucose tolerance[19]. Therefore, single nucleotide polymorphisms (SNPs) of *PTPN1* may influence the development of obesity.

The Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) study (www.helenastudy.com) was designed to provide reliable data on nutrition and health-related phenotypes in a relatively large sample of European adolescents from nine different countries. This cohort includes information on polymorphisms of the *PTPN1* gene, as well as adiposity markers, and biochemical and clinical parameters related to obesity. Therefore, the aim of this study was to assess the association between 5 *PTPN1* polymorphisms and obesity-related phenotypes in European adolescents. We also examined the interaction between physical activity and *PTPN1* polymorphisms on obesity phenotypes to better understand whether physical activity may modulate the genetic predisposition to develop obesity.

Material and methods

Participants

The Healthy Lifestyle in Europe by Nutrition in Adolescence Cross-Sectional Study (HELENA-CSS) attempted to report the lifestyle and nutritional status of European adolescents. A total of 3865 participants (12-18 year-old) from 10 cities in nine European countries (Austria, Belgium, France, Germany, Greece, Hungary, Italy, Spain, and Sweden) were selected to be part of this study. They were randomly selected from public and private schools in each city between October 2006 and December 2007. Multiple measurements were performed by the HELENA Study Group. Blood samples of one-third of these participants (N=1155) were collected with the consequent genetic analysis and clinical biochemistry assays. Among these participants, 1057 individuals (552 girls) with data on *PTPN1* gene polymorphisms along with adiposity markers, and biochemical and clinical parameters related to obesity were included in this study. From that sample, 698 individuals also included data about objectively measured physical activity. Adolescents and corresponding parents/guardians were fully informed about aims and methods of the study such as inclusion criteria[20,21], and signed an informed written consent. The study was conducted according to the Ethical guidelines of the Declaration of Helsinki 1964 (revision of Edinburgh 2000), Good Clinical Practice, and legislation about clinical research in humans in each of the participating countries. The Ethics Committees of each center involved approved the protocol[22].

Assessment of adiposity

Weight and height were measured following standard methods[23]: weight was measured to the nearest 0.1 kg by an electronic scale (SECA 861; Seca Deutschland, Hamburg, Germany) and height to the nearest 0.1 cm by a stadiometer (SECA 225). Waist and hip

circumference were measured in triplicate with an anthropometric inelastic tape (SECA 200; Seca Deutschland, Hamburg, Germany). Thereafter, we calculated waist to height and waist to hip ratios, which were used as surrogate measures of central body fat. Body mass index (BMI) was calculated by dividing weight in kilograms by height in square meters. Using BMI (kg/m^2), we classified individuals as normal weight, overweight or obesity status according to Cole *et al.*[24]. Skinfold thickness was measured to the nearest 0.2 mm in triplicate on the left side of the biceps, triceps, subscapularis, suprailium, thigh, and medial calf with a Holtain Caliper (Holtain Ltd, Crymmych, Wales). Body fat percentage was calculated based on skinfold thicknesses (triceps and subscapular) using the equations by Slaughter *et al.*[25]. Finally, fat mass index (FMI) was calculated by dividing fat mass in kilograms by height in square meters.

Assessment of biochemical and clinical parameters

A total of 30 ml of venous blood was extracted between 8 and 10 am in fasting conditions (ten hours after last meal). Stability of samples regarding transportation between centres had been tested before[26]. Samples were collected in heparinized and gel tubes for serum, maintained in ice and centrifuged (3.500rpm/15min) within 30 min. After centrifugation they were stored and transported (4-7ºC) to the central laboratory (Bonn, Germany), where they were analyzed directly or after being kept at -80ºC. Serum concentrations of biochemical variables were measured in centralized laboratories.

Total cholesterol (TC), high-density cholesterol (HDL-C), low-density cholesterol (LDL-C), apolipoprotein (apo) A1, apoB, triglycerides (TG) and glucose were analyzed on the Dimension RxL clinical chemistry system (Dade Behring, Schwalbach, Germany) with enzymatic methods. Insulin was measured by a solid-phase two-site chemiluminescent immunometric assay with an Immulite 2000 analyzer (DPC Biermann

GmbH, Bad Nauheim, Germany). We computed the homeostatic model assessment (HOMA) as a marker of insulin resistance ([glycaemia X insulin]/22.5). In addition, the quantitative insulin sensitivity check index (QUICKI) was calculated as 1/[log(insulin) + log(glycaemia)]. Leptin concentrations in duplicate were measured using the RayBio® Human Leptin ELISA (Enzyme-Linked Immunosorbent Assay).

Finally, we measured two clinical parameters: systolic and diastolic blood pressure. Blood pressure was measured on the right arm in an extended position with an automatic oscillometric device (OMRON M6). Adolescents quietly sat for 5 min before the measurements. Two measures were taken 5 min apart, and the mean of both values (in mmHg) was used in analyses.

Assessment of physical activity

Physical activity was assessed during seven consecutive days with a uniaxial accelerometer (GT1M; ActiGraph, Pensacola, Florida) attached to the lower back[12]. Adolescents were instructed to wear the accelerometer during all time awake and to remove it only during water-based activities. At least 3 days of recording with a minimum of 8 hours registered per day was set as an inclusion criterion[12]. The time-sampling interval (epoch) was set at 15 seconds. We calculated the time engaged in at least moderate physical activity (≥3 metabolic equivalents) based on a standardized cut-off of 2000 counts/min or more[12]. Moderate to vigorous physical activity (MVPA) was dichotomized into less than 60 min/day (thereafter called inactive adolescents) and 60 min/day or more[12] (thereafter called active adolescents) on average.

Genotyping

Blood for DNA extraction was collected in EDTA K3 tubes, stored at the Analytical Laboratory at the University of Bonn and then sent to the Genomic Analysis Laboratory at the Institut Pasteur de Lille (Lille, France). DNA was extracted from white blood cells with the Puregene kit (Qiagen, Courtaboeuf, France) and stored at -20 °C. Samples were genotyped by an Illumina System (Illumina, Inc, San Diego, California) and the software used was GoldenGate (Inc, San Francisco, California). In the statistical analyses, we considered those tag SNPs that captured multiple SNPs (R^2 value > 0.8 according to the HapMap database). A high rate of genotyping success was achieved (≥99.9%). 5% of the study sample was systematically analyzed in duplicate and the concordance rate was 100%.

Statistical analysis

Deviations from HWE were determined considering a p-value of 0.05 as threshold. Associations between polymorphisms and obesity-related traits were assessed through generalized linear models. We considered five inheritance models (dominant, recessive, overdominant, codominant, additive). Assuming three genotypes "11", "12" and "22", being "1" the minor and "2" the major alleles, respectively, genotypes were coded as follows: dominant: "11-12" and "22"; recessive: "11" and "12-22"; overdominant: "11-22" and "12"; codominant: "11", "12" and "22"; additive: coded as a numerical variable (2 = "11", 1 = "12", 0 = "22"). These models were used with all polymorphisms except for rs10485614. This polymorphism was analyzed using only a dominant model due to the low number of minor homozygotes (Minor allele frequency [MAF] < 0.1; see below). Adjustment variables were age, sex and center. In the case of biochemical and clinical variables, BMI was also included as a covariable in the models. We also considered interactions between SNPs and physical activity with the same models but including an

interaction term (gene*physical activity). For each polymorphism, p-values were computed using a likelihood ratio test between a model with the polymorphism/interaction and a null model without it. Associations in which any level (genotype or genotype*physical activity combination) had lower than 10 samples were discarded. These analyses were performed in the R environment[27] using generalized linear models and assuming a Gaussian distribution. The response variables (i.e., phenotypes) were log-transformed when the assumption of residuals normality was violated. In the case of obesity/overweight, we used a logistic regression assuming a Binomial distribution. We considered the associations between all SNPs and each phenotype under a given inheritance model as the family test, i.e., the number of tests was equal to the number of SNPs analyzed. We selected those genotypes and phenotypes with certain support for their associations to perform interaction and haplotype analyses. In other words, only polymorphisms and phenotypes whose association had certain support were considered for next analyses. Given the exploratory nature of these analyses and the reduced number of independent tests (markers are in linkage disequilibrium, see below), the Bonferroni correction could be too stringent[28]. Instead of this method, we performed an exploratory selection of associations using an approach that controls the expected proportion of false positives (False discovery rate [FDR])[29,30]. We corrected p-values of associations using this approach and then selected those with FDR<0.1 for the next analyses.

Linkage disequilibrium between polymorphisms and haplotype block structures were evaluated with Haploview 4.2 (http://www.broad.mit.edu/mpg/haploview), along with the "haplo.stats"[31] and "SNPassoc"[32] R packages. Haplotype blocks were generated by the algorithm of four-gamete rules[33]. For each block, we tested whether the haplotypes frequencies were deviated from the expectation under linkage equilibrium. Next, we

assessed the association between haplotypes and phenotypes using only additive and dominant models given the low frequency of some haplotypes (see below). Haplotype-phenotype associations were assessed using a permutation procedure with haplo.stat[31]. This approach generates data under the null hypothesis, i.e., no association between obesity-related traits and haplotypes. In each iteration, the values of the phenotype and the covariates are randomly exchanged across individuals, i.e., they are permutated. Given that the phenotype and covariates are shuffled across individuals, but not the genetic data, any signal of association between haplotypes and the phenotype is removed. In other words, we created datasets where there is no association between *PTPN1* haplotypes and obesity-related traits, obtaining a null-distribution. From these datasets, we obtained the expected level of association just by chance according to the structure of our data. If the null hypothesis is true (i.e., no association), changing the combinations of genetic and phenotypic data would not influence the results, i.e., actual and permutated (null) results would be similar. Therefore, the actual results can be compared with a proper null distribution, reducing in that way the risk of false positives. The comparison between the actual data and all permutations (at least 10,000) gave a permutation p-value that was used to assess the significance. Therefore, this is a metric about the strength of the evidence in favor of the null hypothesis of no association between haplotypes and phenotypes[34]. We considered a p-value<0.05 as enough evidence to support a haplotype-phenotype association. We further analyzed these associations to assess what specific haplotype was driving the association along with its direction (i.e., association with higher or lower phenotypic values). We performed linear regressions between the phenotype and the haplotype levels, which were considered as a categorical variable. The most frequent haplotype was used as reference, being the rest of haplotypes compared with it. For each phenotype, several p-values were obtained resulting from the comparison between each

haplotype and the reference haplotype, being these p-values used to calculate the false discovery rate. We considered FDR<0.05 as enough evidence to support the existence of differences between the reference and other haplotypes.

Finally, the interaction analyses between *PTPN1* polymorphisms and physical activity were also corrected for multiple comparison. For each phenotype, the p-values obtained for the interaction between each individual polymorphism and physical activity were corrected with the FDR. We considered a false discovery rate<0.05 as enough support against the null hypothesis of no interaction.

Results

Polymorphisms and characteristics of the sample

Each polymorphism respected the Hardy-Weinberg equilibrium (HWE; P>0.05 in all cases; see Table 1 for allele frequencies and HWE results). Several polymorphisms of the *PTPN1* gene showed linkage disequilibrium between them (Supplementary Figure S1; see Supplementary Table S1 for haplotype frequencies). The characteristics of the study sample are shown in Tables 2 and 3. Characteristics of the specific subset of 698 individuals with physical activity data can be found in Supplementary Tables S2, S3 and S4. These individuals did not differ from the whole cohort in terms of *PTPN1* allele frequencies and obesity-related traits. In addition, *PTPN1* polymorphisms showed the same power to explain phenotypic variability (R^2) in both datasets (p-value=0.49 for a Kruskal-Wallis test between groups). Therefore, we found no evidence of sampling bias within the subset of individuals with physical activity data.

Association between PTPN1 polymorphisms and obesity phenotypes

The exploratory analysis suggested that the five polymorphisms analyzed were associated with adiposity markers, along with biochemical and clinical parameters related to obesity (rs2143511, rs6020608, rs6067472, rs968701, rs10485614). We observed that the major T allele of the rs2143511 polymorphism was associated with lower percentage of obesity/overweight, hip circumference, BMI, body fat percentage and FMI (Table 4). Regarding the rs6020608 polymorphism, we observed an association of the minor T allele with lower hip circumference, BMI, body fat percentage, and FMI. Similarly, the minor T allele of rs6067472 polymorphism was associated with lower hip circumference, BMI, body fat percentage and FMI, along with lower systolic blood pressure. We also found an association between the major G allele of rs968701 polymorphism and lower body fat

percentage and FMI (Table 4). Finally, we observed an association between the rs10485614 polymorphism and leptin, but only under the dominant inheritance model. The AC-CC genotype of rs10485614 was associated with higher leptin levels (p-value=0.02; FDR=0.08; R^2=0.61%; dominant model; see Supplementary Data S1 for all associations analyzed).

Association between PTPN1 polymorphism haplotypes and obesity phenotypes

All *PTPN1* polymorphisms were allocated to a single haplotype block, which showed 7 haplotypes (Supplementary Table S1; Supplementary Figure S1). Haplotype TATTG of *PTPN1* (rs6067472, rs10485614, rs2143511, rs6020608, rs968701; freq=0.279) was associated with lower body fat percentage under additive model (permutation p-value=0.0144) compared to AACCA (freq=0.424; difference between groups=0.05; 95CI=0.09-0.02; p-value=0.0054; FDR=0.0272; differences between groups obtained from models with the response variable log-transformed). The same pattern was found for FMI under additive model, being TATTG haplotype associated with a lower FMI (permutation p-value=0.015) compared to AACCA (difference between groups=0.07; 95CI=0.13-0.02; p-value=0.0046; FDR=0.0228; differences between groups obtained from models with the response variable log-transformed).

Interaction between PTPN1 polymorphisms, physical activity and obesity phenotypes

We found that alleles of the rs2143511, rs6020608, rs6067472 and rs968701 polymorphisms were associated with lower adiposity only in active adolescents (i.e., adolescents spending ≥60 min/day of moderate to vigorous physical activity). Physically active adolescents carrying the major T allele of rs2143511 showed lower hip circumference, BMI, body fat percentage, and FMI (Table 5). Regarding the rs6020608

and rs6067472 polymorphisms, physically active adolescents carrying the minor T alleles of these polymorphisms exhibited lower hip circumference and BMI. Finally, physically active adolescents who were carriers of the major G allele of rs968701 also had lower levels of body fat percentage and FMI (Table 5; see Supplementary Data S2 for all interactions analyzed).

Discussion

The initial and exploratory analyses of the present study showed associations for the T, A, T, T and G alleles of the rs6067472, rs10485614, rs2143511, rs6020608 and rs968701 polymorphisms, respectively, with lower obesity/overweight prevalence and lower values of obesity-related phenotypes (i.e., hip circumference, BMI, body fat percentage, FMI, SBP and leptin). In addition, the haplotype block TATTG constituted by these alleles was associated with lower body fat percentage and FMI. Finally, physical activity may modulate the influence of *PTPN1* polymorphisms on adiposity markers in European adolescents. The T, T, T and G alleles of the rs6067472, rs2143511, rs6020608 and rs968701 polymorphisms, respectively, were associated with lower adiposity in adolescents meeting the physical activity recommendations.

Previous genetic-association studies have linked *PTPN1* polymorphisms with obesity-related phenotypes[19], including type 2 diabetes status, total and low-density cholesterol, glucose metabolism, blood pressure and adiposity markers across several Caucasian and Hispanic cohorts[35–39]. In contrast, other studies have failed to find associations between *PTPN1* polymorphisms and obesity-related traits[19] like glucose regulation, diabetic status or adiposity in variety of cohorts[40,41] including Iranian [42] and Pima populations[43]. With respect to the polymorphisms included in the present study, the rs968701 and rs6020608 have been previously associated with obesity-related phenotypes (glucose regulation and blood pressure), but without considering multiple test correction[35–37]. In contrast, we found these and other *PTPN1* polymorphisms associated with several adiposity markers. Finally, to our knowledge, there is no evidence of association for *PTPN1* and obesity-related traits in Genome Wide Association Studies (GWAS)[44]. This inconsistency can be caused by differences in population-dependent penetrance and allele frequencies, as well as subject age, ethnicity and sample size. In the

case of GWAS, many more markers are analyzed compared to candidate-gene studies, increasing the risk of false positives and making even more important to control for this type of error. However, some methods for multiple-comparison correction can be too conservative [28,30], especially when many genetic markers are analyzed and many of them are in linkage disequilibrium as in the case of GWAS[45]. In that scenario, a stringent control of false positives may lead to the no detection of genes with smaller effects compared to genes strongly associated with obesity-related traits (e.g., *FTO*). To our knowledge, there is no previous evidence about the interaction between physical activity and *PTPN1* polymorphisms in genetic association studies.

PTPN1 is a negative regulator of the insulin signaling pathway[46] as supported by studies showing better glucose homeostasis[47] and lower weight gain[15] in *Ptpn1* knockout mice. Despite the relevance of PTPN1 on insulin signaling, we did not find any association between *PTPN1* polymorphisms and phenotypes related to glucose homeostasis. This could be explained by the differential effect of PTPN1 across tissues. It has been shown that mice with specific *Ptpn1* deletion in muscle had improved glucose homeostasis but no differences in weight gain[47]. In contrast, insulin sensitivity did not increase in adipose tissue of whole-body *Ptpn1* knockout mice despite the decrease of adiposity[18,48]. Therefore, PTPN1 protein could have a different function in adipose tissue. Importantly, the associations found between *PTPN1* polymorphisms and adiposity in the present study could be explained by the influence of PTPN1 on energy balance, as it has been reported increased resting metabolic rate and energy expenditure in *Ptpn1* knockout mice[18]. The impact of PTPN1 on energy balance could be exerted through leptin, given its role in the downregulation of leptin signaling[16,17,49]. Leptin regulates energy homeostasis by influencing food intake and energy expenditure[50]. Accordingly, Bence *et al.*[51] showed increased leptin levels along with reduced food intake and weight in mice

with a neuronal-specific *Ptpn1* deletion. Our results partially support the influence of *PTPN1* on adiposity through leptin, as the AC-CC genotype of rs10485614 was associated with higher leptin levels. In addition, two of the adiposity markers more associated with *PTPN1* polymorphisms, fat mass index and body fat percentage, showed a strong positive correlation with leptin ($P<0.0001$ and $\rho>0.6$ in both cases; Supplementary Figure S2). These results suggest that leptin resistance could have a relevant role in the patterns of adiposity in this cohort and may be mediated by *PTPN1*. However, we must bear in mind that rs10485614 was only associated with leptin according to FDR<0.1 and was not associated with adiposity. This could be caused by the lower number of minor homozygotes for this polymorphism in our cohort, preventing the test of associations under non-dominant models.

PTPN1 could influence energy expenditure not only through leptin, but also through leptin-independent pathways. Several studies have shown increased energy expenditure and decreased adiposity in *Ptpn1* knockout mice without differences in food intake[18,52–54]. Interestingly, other studies reported lower weight despite existing no differences in food intake for *Ptpn1* knockout mice lacking leptin-responsive hypothalamic neurons[16] or with leptin deletion[17], compared to mice having active *Ptpn1*. These results suggest a role of *PTPN1* on energy expenditure and adiposity independently of leptin in general, and leptin-reduced satiety in particular. The pathways (leptin-dependent or independent) through which PTPN1 favor energy expenditure could implicate the activation of brown adipose tissue (BAT), a relevant tissue on non-shivering thermogenesis[55]. In this regard, it has been shown higher activity of AMP-activated protein kinase (AMPK), a mediator of leptin's effects[56], in BAT and muscle tissue of *Ptpn1* knockout mice, leading to gene expression changes that enhanced energy expenditure[57]. Similarly, the deletion of *Ptpn1* in mice seems to favour BAT mass and

activity after cold exposure[58], along with increase brown adipogenesis[59,60] (but see Klaman *et al*.[18] for contrary results). Finally, the influence of PTPN1 on energy homeostasis and adiposity could be mediated by the tropomyosin receptor kinase B (TrkB) and its ligand, the brain-derived neurotrophic factor (BDNF), as both play a key role in the regulation of energy homeostasis[61]. It has been shown that *Ptpn1* knockout mice have higher phosphorylation of TrkB and higher increases of core temperature under a treatment of BDNF compared to wildtypes[62]. This suggests that the deletion of *Ptpn1* can enhance BDNF-mediated energy expenditure, supporting that PTPN1 could influence energy homeostasis, and consequently adiposity, through different pathways.

Finally, the association of the minor T allele of rs6067472 with lower adiposity and SBP is congruent with results in mice showing a protective effect of *PTPN1* deletion against adrenergic hypertension. Bruder-Nascimento *et al*.[54] reported that *Ptpn1* knockout mice suffered lower increases of arterial pressure under chronic sympatho-activation. Note, however, that we did not find a strong signal for this association in our study. SBP was associated with *PTPN1* polymorphisms only in the first, exploratory analysis (FDR<0.1), without enough support for the association in further analyses.

We not only found that *PTPN1* polymorphisms are associated with obesity-related traits, but also that their influence may be modulated by physical activity. When considering the interaction with physical activity, *PTPN1* alleles were associated with less adiposity only in physically active adolescents. These results suggest the existence of a synergy between *PTPN1* and physical activity. Evidence coming from mice suggests that physical activity could decrease the content of PTPN1[63,64], while Guerra *et al*.[65] showed an increase of PTPN1 levels in the vastus lateralis of human subjects under bed rest. In addition, it has been shown in mice that aging increases levels of liver PTPN1, adiposity, and the inhibition of insulin signaling[63]. All these changes associated with

aging were reverted when mice were exposed to acute exercise. Therefore, *PTPN1* variants associated with less content/functionality of PTPN1 in combination with the attenuation of PTPN1 effects by high levels of physical activity may favor a greater reduction of adiposity, that is, they may have a synergistic effect. There are several mechanisms through which the interaction between physical activity and PTPN1 could occur. For example, it could be mediated by Sirtuin 1 (SIRT1). SIRT1 positively impacts insulin resistance through the repression of *PTPN1* in cell cultures[66], being upregulated and having the same effect in the skeletal muscle of old mice exposed to exercise[67]. Therefore, physical activity could inhibit PTPN1 through SIRT1 upregulation. However, we did not find enough support for the association between markers of glucose regulation and *PTPN1* polymorphisms. An alternative explanation for the interaction between *PTPN1* polymorphisms and physical activity on adiposity could be related to energy expenditure and homeostasis. As previously mentioned, PTPN1 deletion seems to increase mass and activity of BAT and upregulate the BDNF pathway. Interestingly, physical activity could also increase energy expenditure through similar routes[68,69]. This is congruent with our results showing that *PTPN1* polymorphisms were associated with lower adiposity only in physically active adolescents. High levels of physical activity in combination with genetic variants that reduce the functionality or content of PTPN1 may have a synergistic effect on energy expenditure, decreasing adiposity.

All the *PTPN1* variants analyzed are intronic, but they could exert an effect on PTPN1 through gene expression. Data from GTEx (https://gtexportal.org/) shows contrasting results for the polymorphisms included in the present study. For example, the rs10485614 polymorphism is associated with lower gene expression for the AA genotype respect to AC[70]. In our study, the AC-CC genotype of rs10485614 was associated with higher leptin levels (FDR=0.08), which is congruent with the hypothesis that alleles

associated with lower PTPN1 levels may be protective regarding obesity-related traits. However, other variants showed opposite results, like the rs2143511 polymorphism. The protective T allele of this polymorphism is associated with higher *PTPN1* expression[71]. Note, however, that these results come from whole blood, not adipose or muscle tissue. As previously mentioned, the effect of PTPN1 could vary depending on the tissue. We have not found different expression levels associated with the studied polymorphisms for adipose and muscle tissues, which are the most relevant tissues for interpreting our results. Another possible explanation is that these polymorphisms are in linkage with a coding variant influencing the function of PTPN1. We found certain support for this hypothesis when comparing the explicative power between several models: i) Model including the *PTPN1* haplotype; ii) Model including the 5 *PTPN1* polymorphisms as independent predictors; iii) Models lacking one of the polymorphisms each time. Each individual variant showed little independent explicative power (Supplementary Data S3). In other words, each individual polymorphism does not explain much phenotypic variability independently of the rest of variants (further details also explained in supplementary appendix). This suggests that the polymorphisms included in this study are possibly tagging a unique causal variant that could influence the function of PTPN1. Given the lack of data about PTPN1 levels for specific tissues in the present study or cell culture experiments for PTPN1 activity with the studied variants, possible mechanisms of action can only be hypothesized. Therefore, more research is needed to elucidate the specific mechanisms by which these *PTPN1* variants and physical activity may influence adiposity and other health markers.

Limitations of the present study should be considered. Firstly, this is a cross-sectional study, hence cause-effect relationships cannot be determined. Our results should be tested in future experimental studies to assess direct causal correlation between *PTPN1*

polymorphisms and obesity-related phenotypes, along with the interaction with physical activity. In addition, these associations could be modified by gene-gene and other gene-environmental interactions. Finally, we have no information about patterns of relatedness between participants, and we do not know the ethnic/racial origin of the sample. Our results should be considered carefully and studies with larger sample size could help to further support the role of *PTPN1* gene on obesity and its interaction with physical activity.

In summary, we found that *PTPN1* polymorphisms are associated with adiposity markers in European adolescents. Our results also suggest that the influence of rs6067472, rs2143511, rs6020608 and rs968701 on adiposity markers can be modulated by physical activity. *PTPN1* polymorphisms were only associated with lower adiposity in physically active individuals. Therefore, those individuals who meet the recommendations of daily physical activity may have a potential benefit, that is, a reduction of obesity risk by modulating the genetic predisposition to obesity.

Data availability statement

In order to maintain a low risk of re-identification and compliance with the EU General Data Protection Regulation, the clinical and genetic raw data used in this study cannot be public. Interested parties can request the Risk Analysis Assessment approved by the funding agency that includes full explanations about the security measures and their justification.

References

1. GBD 2015 Obesity Collaborators. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *N. Engl. J. Med.* **377**, 13–27 (2017).
2. Xu, B., & Xie, X. Neurotrophic Factor Control of Satiety and Body Weight. *Nat Rev Neurosci* **17**, 282–92 (2016).
3. Abdalla, M.M.I. Central and peripheral control of food intake. *Endocr. Regul.* **51**, 52–70 (2017).
4. Loos, R.J.F., & Yeo, G.S.H. The genetics of obesity: from discovery to biology. *Nat. Rev. Genet.* **23**, 120–33 (2022).
5. Farooqi, I.S., & O'Rahilly, S. Genetics of obesity in humans. *Endocr. Rev.* **27**, 710–8 (2006).
6. Sheikh, A.B. et al. The Interplay of Genetics and Environmental Factors in the Development of Obesity. *Cureus* **9**, e1435 (2017).
7. Ruiz, J.R. et al. Attenuation of the effect of the FTO rs9939609 polymorphism on total and central body fat by physical activity in adolescents: The HELENA study. *Arch. Pediatr. Adolesc. Med.* **164**, 328–33 (2010).
8. Pascual-Gamarra, J.M. et al. Association between *UCP1*, *UCP2*, and *UCP3* gene polymorphisms with markers of adiposity in European adolescents: The HELENA study. *Pediatr. Obes.* **14**, e12504 (2019).
9. Tanisawa, K., Tanaka, M., & Higuchi, M. Gene-exercise interactions in the development of cardiometabolic diseases. *J. Phys. Fit. Sport. Med.* [Internet] **5**, 25–36 (2016). Available from: https://doi.org/10.7600/jpfsm.5.25
10. Smith, J.K. Exercise and Atherogenesis. *Exerc. Sport Sci. Rev.* **29**, 49–53 (2001).

11. Graff, M. et al. Genome-wide physical activity interactions in adiposity — A meta-analysis of 200,452 adults. *PLoS Genet.* **13**, 1–26 (2017).
12. Ruiz, J.R. et al. Objectively measured physical activity and sedentary time in European adolescents. *Am. J. Epidemiol.* **174**, 173–84 (2011).
13. Salazar-Tortosa, D.F. et al. Association between lipoprotein lipase gene polymorphisms and cardiovascular disease risk factors in European adolescents: The Healthy Lifestyle in Europe by Nutrition in Adolescence study. *Pediatr. Diabetes* **21**, 747–757 (2020).
14. Cho, H. Protein Tyrosine Phosphatase 1B (PTP1B) And Obesity. *Vitam. Horm.* [Internet] **91**, 405–24 (2013). Available from: http://dx.doi.org/10.1016/B978-0-12-407766-9.00017-1
15. Elchebly, M. et al. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science (80-.).* **283**, 1544–8 (1999).
16. Zabolotny, J.M. et al. PTP1B regulates leptin signal transduction in vivo. *Dev. Cell* **2**, 489–95 (2002).
17. Cheng, A. et al. Attenuation of leptin action and regulation of obesity by protein tyrosine phosphatase 1B. *Dev. Cell* **2**, 497–503 (2002).
18. Klaman, L.D. et al. Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in Protein-Tyrosine Phosphatase 1B-deficient mice. *Mol. Cell. Biol.* **20**, 5479–89 (2000).
19. Tsou, R.C., & Bence, K.K. The genetics of *PTPN1* and obesity: Insights from mouse models of tissue-specific PTP1B deficiency. *J. Obes.* **2012**, 1–8 (2012).
20. Moreno, L.A. et al. Assessing, understanding and modifying nutritional status, eating habits and physical activity in European adolescents: the HELENA

(Healthy Lifestyle in Europe by Nutrition in Adolescence) Study. *Public Health Nutr.* **11**, 288–99 (2007).

21. Moreno, L.A. et al. Design and implementation of the Healthy Lifestyle in Europe by Nutrition in Adolescence Cross-Sectional Study. *Int. J. Obes.* **32**, S4–11 (2008).
22. Béghin, L. et al. Quality assurance of ethical issues and regulatory aspects relating to good clinical practices in the HELENA Cross-Sectional Study. *Int. J. Obes.* **32 Suppl 5**, S12–8 (2008).
23. Nagy, E. et al. Harmonization process and reliability assessment of anthropometric measurements in a multicenter study in adolescents. *Int. J. Obes.* **32**, S58-65 (2008).
24. Cole, T.J., Bellizzi, M.C., Flegal, K.M., & Dietz, W.H. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* **320**, 1240–3 (2000).
25. Slaughter, M. et al. Skinfold equations for estimation of body fatness in children and youth. *Hum. Biol.* **60**, 709–23 (1998).
26. González-Gross, M. et al. Sampling and processing of fresh blood samples within a European multicenter nutritional study: Evaluation of biomarker stability during transport and storage. *Int. J. Obes.* **32**, S66–75 (2008).
27. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. (2017).
28. Sham, P.C., & Purcell, S.M. Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* **15**, 335–46 (2014).
29. Benjamini, Y., & Hochberg, Y. Controlling the False Discovery Rate: A

Practical and Powerful Approach to Multiple Testing. *J. Royal Stat. Soc.* **57**, 289–300 (1995).

30. Qu, H.-Q., Tien, M., & Polychronakos, C. Statistical significance in genetic association studies. *Clin. Investig. Med.* **33**, E266–70 (2010).

31. Sinnwell, J.P., & Schaid, D.J. haplo.stats: Statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous [Internet]. (2016). Available from: https://cran.r-project.org/package=haplo.stats

32. González, J.R. et al. SNPassoc: An R package to perform whole genome association studies. *Bioinformatics* **23**, 644–5 (2007).

33. Wang, N. et al. Distribution of recombination crossovers and the origin of haplotype blocks: The interplay of population history, recombination, and mutation. *Am. J. Hum. Genet.* **71**, 1227–34 (2002).

34. Sterne, J.A.C., & Smith, G.D. Sifting the evidence-what's wrong with significance tests? *BMJ* [Internet] **322**, 226–31 (2001). Available from: https://www.bmj.com/content/322/7280/226.1

35. Palmer, N.D. et al. Association of protein tyrosine phosphatase 1B gene polymorphisms with measures of glucose homeostasis in Hispanic Americans: The Insulin Resistance Atherosclerosis Study (IRAS) family study. *Diabetes* **53**, 3013–9 (2004).

36. Spencer-Jones, N.J. et al. Protein tyrosine phosphatase-1B gene PTPN1: Selection of tagging single nucleotide polymorphisms and association with body fat, insulin sensitivity, and the metabolic syndrome in a normal female population. *Diabetes* **54**, 3296–304 (2005).

37. Cheyssac, C. et al. Analysis of common *PTPN1* gene variants in type 2 diabetes, obesity and associated phenotypes in the French population. *BMC*

Med. Genet. **7**, 1–10 (2006).

38. Bento, J.L. et al. Association of protein tyrosine phosphatase 1B gene polymorphisms with type 2 diabetes. *Diabetes* **53**, 3007–12 (2004).
39. Bauer, F. et al. *PTPN1* polymorphisms are associated with total and low-density lipoprotein cholesterol. *Eur. J. Prev. Cardiol.* **17**, 28–34 (2010).
40. Florez, J.C. et al. Association testing of the protein tyrosine phosphatase 1B gene (*PTPN1*) with type 2 diabetes in 7,883 people. *Diabetes* **54**, 1884–91 (2005).
41. Bauer, F. et al. No association of *PTPN1* polymorphisms with macronutrient intake and measures of adiposity. *Obesity* **16**, 2767–71 (2008).
42. Meshkani, R. et al. Polymorphisms within the protein tyrosine phosphatase IB (*PTPN1*) gene promoter: Functional characterization and association with type 2 diabetes and related metabolic traits. *Clin. Chem.* **53**, 1585–92 (2007).
43. Traurig, M. et al. Protein tyrosine phosphatase 1B is not a major susceptibility gene for type 2 diabetes mellitus or obesity among Pima Indians. *Diabetologia* **50**, 985–9 (2007).
44. EMBL-EBI. GWAS Catalogue: PTPN1 [Internet]. (2022). Available from: https://www.ebi.ac.uk/gwas/genes/PTPN1
45. The International HapMap Consortium. A haplotype map of the human genome. *Nature* [Internet] **437**, 1299–320 (2005). Available from: https://doi.org/10.1038/nature04226
46. Geraldes, P. Protein phosphatases and podocyte function. *Curr. Opin. Nephrol. Hypertens.* **27**, 49–55 (2018).
47. Delibegovic, M. et al. Improved Glucose Homeostasis in Mice with Muscle-Specific Deletion of Protein-Tyrosine Phosphatase 1B. *Mol. Cell. Biol.* **27**,

7727–34 (2007).

48. Ramachandran, C., & Kennedy, B.P. Protein Tyrosine Phosphatase 1B: A Novel Target for Type 2 Diabetes and Obesity. *Curr. Top. Med. Chem.* **3**, 749–57 (2003).

49. White, C.L. et al. HF diets increase hypothalamic PTP1B and induce leptin resistance through both leptin-dependent and -independent mechanisms. *Am. J. Physiol. - Endocrinol. Metab.* **296**, E291–E299 (2009).

50. Ahima, R.S., & Flier, J.S. Leptin. *Annu. Rev. Physiol.* **62**, 413–37 (2000).

51. Bence, K.K. et al. Neuronal PTP1B regulates body weight, adiposity and leptin action. *Nat. Med.* **12**, 917–24 (2006).

52. Banno, R. et al. PTP1B and SHP2 in POMC neurons reciprocally regulate energy balance in mice. *J. Clin. Invest.* **120**, 720–34 (2010).

53. Bruder-Nascimento, T. et al. Ptp1b deletion in pro-opiomelanocortin neurons increases energy expenditure and impairs endothelial function via TNF-α dependent mechanisms. *Clin. Sci.* **130**, 881–93 (2016).

54. Bruder-Nascimento, T. et al. Deletion of protein tyrosine phosphatase 1b in proopiomelanocortin neurons reduces neurogenic control of blood pressure and protects mice from leptin- and sympatho-mediated hypertension. *Pharmacol. Res.* **102**, 235–44 (2015).

55. Cannon, B., & Nedergaard, J. Brown Adipose Tissue: Function and Physiological Significance. *Physiol. Rev.* [Internet] **84**, 277–359 (2004). Available from: http://physrev.physiology.org/cgi/doi/10.1152/physrev.00015.2003

56. Minokoshi, Y. et al. Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* **415**, 339–43 (2002).

57. Xue, B. et al. Neuronal Protein Tyrosine Phosphatase 1B deficiency results in Inhibition of hypothalamic AMPK and isoform-specific activation of AMPK in peripheral tissues. *Mol. Cell. Biol.* **29**, 4563–73 (2009).

58. de Jonghe, B.C. et al. Deficiency of PTP1B in POMC neurons leads to alterations in energy balance and homeostatic response to cold exposure. *Am. J. Physiol. - Endocrinol. Metab.* **300**, E1002–E1011 (2011).

59. Miranda, S. et al. Beneficial effects of PTP1B deficiency on brown adipocyte differentiation and protection against apoptosis induced by pro- and anti-inflammatory stimuli. *Cell. Signal.* **22**, 645–59 (2010).

60. Matsuo, K. et al. Regulation of brown fat adipogenesis by protein tyrosine phosphatase 1B. *PLoS One* **6** (2011).

61. Vanevski, F., & Xu, B. Molecular and neural bases underlying roles of BDNF in the control of body weight. *Front. Neurosci.* **7**, 1–10 (2013).

62. Ozek, C. et al. Protein-tyrosine phosphatase 1B (PTP1B) is a novel regulator of central brain-derived neurotrophic factor and tropomyosin receptor kinase B (TrkB) signaling. *J. Biol. Chem.* **289**, 31682–92 (2014).

63. de Moura, L.P. et al. Acute exercise decreases PTP-1B protein level and improves insulin signaling in the liver of old rats. *Immun. Ageing* **10**, 1–9 (2013).

64. Ropelle, E.R. et al. Reversal of diet-induced insulin resistance with a single bout of exercise in the rat: The role of PTP1B and IRS-1 serine phosphorylation. *J. Physiol.* **577**, 997–1007 (2006).

65. Guerra, B. et al. Leptin signaling in skeletal muscle after bed rest in healthy humans. *Eur. J. Appl. Physiol.* **114**, 345–57 (2014).

66. Sun, C. et al. SIRT1 improves insulin sensitivity under insulin-resistant

conditions by repressing PTP1B. *Cell Metab.* **6**, 307–19 (2007).

67. Pauli, J.R. et al. Acute exercise reverses aged-induced impairments in insulin signaling in rodent skeletal muscle. *Mech. Ageing Dev.* **131**, 323–9 (2010).

68. Dinoff, A., Herrmann, N., Swardfager, W., & Lanctôt, K.L. The effect of acute exercise on blood concentrations of brain-derived neurotrophic factor (BDNF) in healthy adults: A meta-analysis. *Eur J Neurosci* **46**, 1635–46 (2017).

69. Valgas da Silva, C.P., Hernández-Saavedra, D., White, J.D., & Stanford, K.I. Cold and Exercise : Therapeutic Tools to Activate Brown Adipose Tissue and Combat Obesity. *Biology (Basel).* **8**, 1–29 (2019).

70. GTEx Consortium. GTEx v8 eQTLs: rs10485614 PTPN1 [Internet]. Available from: https://www.gtexportal.org/home/snp/rs10485614

71. GTEx Consortium. GTEx v8 eQTLs: rs2143511 PTPN1 [Internet]. Available from: https://www.gtexportal.org/home/snp/rs2143511

Acknowledgements

We wish to thank Rosa Maria Torres, Petra Pickert, and Anke Berchtold for their contribution to laboratory work

Funding

The HELENA Study was supported by contract FOOD-CT-2005-007034 from the European Community Sixth RTD Framework Programme. The present study was also supported by a Marie S. Curie Global Fellowship within the European Union research and innovation framework programme (2014-2020; ClimAHealth; https://doi.org/10.3030/101030971), by a grant from the Spanish Ministry of Science and Innovation (AGL2007-29784-E), by *Redes Temáticas de Investigación Cooperativa RETIC* grant (Red SAMID RD16/0022), by the University of Granada Plan Propio de Investigación (Excellence actions: Unit of Excellence on Exercise and Health [UCEES]), by the *Junta de Andalucía, Consejería de Economía, Conocimiento, Empresas y Universidad* (refs. P18-624 RT-4455, SOMM17/6107/UGR), and by grants from the Public University of Navarra, "Ayudas a Grupos de Investigación (2019)". The content of this article reflects only the authors' views, and the European Community is not liable for any use that may be made of the information contained therein.

Author contributions

IL, MGG, MSC, LAM, AGZ, KW, AM, DE, DFST and JRR designed the study; DFST performed all analyses; DFST and JRR wrote the initial draft and all co-authors significantly contributed to the final version.

Competing interests

The authors declare that they have no competing interests.

Consent statement

As indicated in the Methods section, adolescents and corresponding parents/guardians were fully informed about aims and methods of the study and signed an informed written consent.

Tables

Table 1: *PTPN1* alleles, p-value for Hardy-Weinberg equilibrium (HWE) according to the present study and minor allele frequencies (MAF). The allele frequencies of the minor alleles[a] according to the HELENA study are shown considering all centers, but also within each center. In addition, the minor allele frequencies in Europeans populations according to the 1000 Genomes Project (1KGP) are shown (https://www.ncbi.nlm.nih.gov/snp/). All alleles are reported in the forward orientation.

					Minor allele frequencies within centers										
	Major allele	**Minor allele**	**HWE**	**All centers**	**Athens Greece**	**Dortmund Germany**	**Gent Belgium**	**Heraklion Crete**	**Lille France**	**Pecs Hungary**	**Roma Italy**	**Stockholm Sweden**	**Vienna Austria**	**Zaragoza Spain**	**1KGP Europe**
rs6067472	A	T	0.61	0.39	0.39	0.37	0.39	0.35	0.39	0.46	0.38	0.43	0.37	0.37	0.36
rs10485614	A	C	0.36	0.07	0.09	0.08	0.07	0.09	0.07	0.03	0.09	0.06	0.08	0.06	0.08
rs2143511	T	C	0.53	0.44	0.44	0.46	0.41	0.49	0.41	0.41	0.42	0.4	0.5	0.45	0.45
rs6020608	C	T	0.6	0.29	0.29	0.28	0.29	0.25	0.26	0.36	0.3	0.31	0.25	0.25	0.27
rs968701	G	A	1	0.5	0.47	0.52	0.48	0.52	0.5	0.48	0.5	0.47	0.55	0.49	0.49

[a]The minor allele was determined in the overall sample (hence some MAF are higher than 0.5).

Table 2: Sample size and obesity/overweight prevalence (i.e., both obesity and overweight status included) by center and sex.

Center	All		Male		Female	
	Sample size	Overweight (%)	Sample size	Overweight (%)	Sample size	Overweight (%)
Athens in Greece	104	27.88	45	24.44	59	30.51
Dortmund in Germany	116	26.72	67	26.87	49	26.53
Gent in Belgium	109	6.42	56	0	53	13.21
Heraklion in Crete	103	37.86	52	46.15	51	29.41
Lille in France	87	18.39	36	19.44	51	17.65
Pecs in Hungary	137	18.98	53	16.98	84	20.24
Roma in Italy	80	38.75	47	42.55	33	33.33
Stockholm in Sweden	100	17	50	22	50	12
Vienna in Austria	111	17.12	48	22.92	63	12.7
Zaragoza in Spain	110	18.18	51	27.45	59	10.17

Table 3: Characteristics of the study population by sex and weight status. Data are expressed as mean ± standard deviation.

Phenotype	All		Male		Female	
	Non-overweight (n=822)	Overweight (n=235)	Non-overweight (n=380)	Overweight (n=125)	Non-overweight (n=442)	Overweight (n=110)
Age (y)	14.75±1.2	14.59±1.29	14.78±1.26	14.62±1.23	14.71±1.15	14.56±1.37
Weight (kg)	54.53±8.8	73.38±13.24	56.76±9.98	77.38±14.26	52.62±7.11	68.83±10.29
Height (cm)	165.41±9.27	165.62±9.61	169.44±9.99	169.68±9.69	161.94±6.94	161.02±7.16
BMI (kg/m^2)	19.83±2.02	26.61±3.29	19.62±1.96	26.76±3.76	20.02±2.04	26.45±2.65
Waist circum. (cm)	69.27±5.61	83.55±8.56	70.68±5.29	85.95±8.77	68.06±5.6	80.89±7.5
Waist/Height ratio	0.42±0.03	0.5±0.05	0.42±0.03	0.51±0.05	0.42±0.03	0.5±0.04
Hip circum. (cm)	88.78±6.65	101.47±8.29	87.06±6.45	100.68±8.17	90.27±6.47	102.35±8.37
Waist/Hip ratio	0.78±0.05	0.83±0.07	0.81±0.04	0.85±0.05	0.75±0.05	0.79±0.08
Body fat (%)	20.41±6.98	34.57±8.9	15.54±5.69	34.3±11.05	24.19±5.36	34.86±5.81
FMI (kg/m^2)	4.12±1.67	9.35±3.25	3.09±1.32	9.38±3.91	4.92±1.46	9.32±2.35
Total cholesterol (mg/dL)	160.26±27.36	162.4±28.82	152.91±24.72	157.45±29.86	166.6±27.96	168.02±26.62
LDL-C (mg/dL)	92.7±24.83	100.72±25.07	88.31±23	98.25±26.68	96.48±25.73	103.53±22.91
HDL-C (mg/dL)	56.63±10.57	50.46±9.59	54.61±10.24	48.8±8.36	58.38±10.55	52.35±10.55
Total cholesterol/HDL-C	2.9±0.61	3.29±0.72	2.87±0.64	3.29±0.74	2.92±0.59	3.3±0.69
LDL-C/HDL-C	1.7±0.59	2.06±0.66	1.69±0.62	2.07±0.68	1.71±0.57	2.06±0.64
Triglycerides (mg/dL)	65.41±31.04	81.52±44.41	59.76±27.5	77.41±39.03	70.27±33.05	86.19±49.6
Triglycerides/HDL-C	1.22±0.74	1.72±1.17	1.17±0.68	1.68±1.09	1.28±0.78	1.76±1.26

ApoA1 (mg/dL)	1.53±0.22	1.43±0.22	1.48±0.21	1.39±0.18	1.57±0.22	1.48±0.25
ApoB (mg/dL)	0.64±0.15	0.7±0.16	0.61±0.14	0.68±0.16	0.67±0.16	0.73±0.16
ApoB/ApoA1	0.43±0.12	0.5±0.14	0.42±0.12	0.5±0.15	0.43±0.12	0.51±0.14
ApoB/LDL-C	0.27±0.03	0.27±0.03	0.27±0.03	0.27±0.02	0.27±0.03	0.27±0.03
Insulin (micro lU/mL)	9.26±6.15	14.04±11.15	8.8±6.29	14.29±13.11	9.67±6.01	13.75±8.29
HOMA	2.1±1.52	3.24±2.87	2.03±1.53	3.39±3.45	2.16±1.51	3.07±1.99
QUICKI	0.35±0.03	0.33±0.03	0.36±0.03	0.33±0.03	0.35±0.03	0.33±0.02
Leptin (ng/ml)	14.93±16.41	36.42±30.63	5.41±6.34	22.97±22.21	22.87±17.93	50.12±32.02
SBP (mm Hg)	118.29±12.52	126.08±14.16	122.14±13.1	130.29±14.63	115±11.01	121.37±12.04
DBP (mm Hg)	67.38±8.62	70.26±9.23	66.6±8.56	70.31±9.39	68.05±8.63	70.21±9.09

BMI: body mass index (calculated as weight in kilograms divided by height in square meters); FMI: fat mass index (calculated as fat mass in kilograms divided by height in square meters); TC: total cholesterol; LDL-C: low-density cholesterol; HDL-C: high-density cholesterol; TG: triglycerides; Apo: apolipoprotein; HOMA: homeostatic model assessment; QUICKI: quantitative insulin sensitivity check index; SBP: systolic blood pressure; DBP: diastolic blood pressure.

Table 4: Associations between the studied polymorphisms and obesity-related phenotypes with FDR<0.1 under additive or codominant models. Mean ± standard deviation is reported for phenotypes in each genotype. For each association, it is reported the p-value, FDR and R^2 under additive and codominant inheritance models. Values are adjusted center, sex, and age (also BMI for biochemical and clinical variables).

SNP	Phenotype	11	12	22	Additive			Codominant		
					P	FDR	R^2	P	FDR	R^2
rs2143511	% obesity/overweight	26.57	23.58	17.65	0.022	0.086	0.499	0.068	0.273	0.507
	Hip circum. (cm)	92.02±8.94	91.97±8.66	90.69±8.85	0.019	0.067	0.535	0.042	0.138	0.616
	BMI (kg/m^2)	21.71±3.68	21.43±3.66	20.99±3.68	0.016	0.064	0.553	0.055	0.217	0.554
	Body fat (%)	24.24±9.5	24±9.42	22.54±9.56	0.01	0.031	0.671	0.031	0.098	0.697
	FMI (kg/m^2)	5.54±3.13	5.4±3.03	4.97±2.99	0.008	0.029	0.705	0.028	0.085	0.719
rs6020608	Hip circum. (cm)	89.66±9.11	91.57±8.6	91.87±8.85	0.033	0.067	0.441	0.069	0.138	0.521
	BMI (kg/m^2)	20.74±3.32	21.29±3.66	21.48±3.74	0.062	0.083	0.331	0.163	0.217	0.347
	Body fat (%)	21.65±9.59	23.3±9.44	24.11±9.5	0.019	0.031	0.553	0.06	0.098	0.564
	FMI (kg/m^2)	4.72±2.92	5.2±2.98	5.45±3.1	0.019	0.029	0.553	0.06	0.085	0.567
rs6067472	Hip circum. (cm)	90.86±8.5	91.37±8.89	92.11±8.75	0.057	0.077	0.352	0.162	0.217	0.355
	BMI (kg/m^2)	21.04±3.36	21.23±3.78	21.62±3.65	0.055	0.083	0.351	0.142	0.217	0.373
	Body fat (%)	22.35±9.38	23.48±9.63	24.19±9.36	0.023	0.031	0.517	0.073	0.098	0.526
	FMI (kg/m^2)	4.94±2.87	5.23±3.08	5.5±3.05	0.022	0.029	0.528	0.064	0.085	0.553
	SBP (mm Hg)	119.07±13.29	119.16±12.71	121.57±13.94	0.021	0.084	0.516	0.056	0.223	0.559
rs968701	Body fat (%)	23.63±9.27	23.86±9.69	22.94±9.37	0.054	0.054	0.375	0.14	0.14	0.396

FMI (kg/m^2)	5.33±2.99	5.39±3.16	5.05±2.85	0.051	0.051	0.383	0.131	0.131	0.409

SNP: single nucleotide polymorphism; 1: minor allele; 2: major allele; P: p-value; FDR: false discovery rate; R^2: explicative power in percentage, see supplementary material (Supplementary Data S1) for details about its calculation; additive: additive inheritance model; codominant: codominant inheritance model; BMI: body mass index (calculated as weight in kilograms divided by height in square meters); FMI: fat mass index (calculated as fat mass in kilograms divided by height in square meters); SBP: systolic blood pressure.

Table 5: Interactions between the studied polymorphisms and physical activity with FDR<0.05 under additive or codominant models. Mean ± standard deviation is reported for phenotypes in each genotype*physical activity level. For each association, it is reported the p-value, FDR and R^2 under additive and codominant inheritance models. Values are adjusted center, sex, and age (also BMI for biochemical and clinical variables).

SNP	Phenotype	11 PI	12 PI	22 PI	11 PA	12 PA	22 PA	Additive			Codominant		
								P	FDR	R^2	P	FDR	R^2
rs2143511	Hip circum. (cm)	90.95±8.78	91.3±8.43	92.15±8.55	90.07±9.04	91.38±8.62	87.95±8.74	0.029	0.039	0.702	0.011	0.014	1.335
	BMI (kg/m^2)	20.9±3.54	20.81±3.32	21.26±3.33	21.17±3.17	21.45±3.36	20.18±3.12	0.015	0.029	0.868	0.005	0.012	1.564
	Body fat (%)	24.06±9.52	24.18±8.58	24.66±9.45	21.61±9.62	22.03±9.46	18.85±7.97	0.014	0.028	0.896	0.034	0.068	1.005
	FMI (kg/m^2)	5.3±3.04	5.26±2.74	5.48±2.96	4.8±2.87	4.98±2.85	3.96±2.27	0.008	0.017	1.032	0.013	0.026	1.287
rs6020608	Hip circum. (cm)	92.62±8.76	91.36±8.31	91.41±8.65	85.57±8.94	90.05±8.62	90.71±8.79	0.069	0.069	0.488	0.016	0.016	1.208
	BMI (kg/m^2)	21.62±2.99	20.84±3.26	20.96±3.48	19.46±2.95	21.01±3.04	21.2±3.46	0.103	0.103	0.387	0.012	0.012	1.278
rs6067472	Hip circum. (cm)	93.22±8.04	90.97±8.34	91.49±8.89	87.16±8.8	90.31±8.9	90.68±8.61	0.024	0.039	0.746	0.003	0.01	1.739
	BMI (kg/m^2)	21.52±3.44	20.74±3.16	21.04±3.58	20.03±3.11	20.97±3.22	21.34±3.39	0.042	0.055	0.605	0.012	0.012	1.281
rs968701	Body fat (%)	23.27±9.04	24.34±8.78	25.13±9.42	22.13±9.73	20.97±9.39	19.33±7.75	0.005	0.021	1.159	0.019	0.068	1.174
	FMI (kg/m^2)	5.05±2.82	5.33±2.84	5.59±2.92	4.94±2.86	4.69±2.8	4.06±2.29	0.003	0.012	1.296	0.013	0.026	1.293

SNP: single nucleotide polymorphism; 1: minor allele; 2: major allele; PI: physically inactive; PA: physically active; P: p-value; FDR: false discovery rate; R^2: explicative power in percentage, see supplementary material (Supplementary Data S2) for details about its calculation; additive: additive inheritance model; codominant: codominant inheritance model; BMI: body mass index (calculated as weight in kilograms divided by height in square meters); FMI: fat mass index (calculated as fat mass in kilograms divided by height in square meters).