Genetic variation at the *CY2C19* gene associated with Metabolic Syndrome susceptibility in a South Portuguese population: results from the pilot study of the European Health Examination Survey in Portugal

Vânia Gaio[1], Baltazar Nunes[1], Aida Fernandes[2], Francisco Mendonça[3], Filomena Horta Correia[3], Álvaro Beleza[2], Ana Paula Gil[1], Mafalda Bourbon[4], Astrid Vicente[4], Carlos Matias Dias[1], Marta Barreto da Silva[1]

[1] Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal; [2] Laboratório de Saúde Pública Dra. Laura Ayres, Faro, Portugal; [3] Administração Regional de Saúde do Algarve, Faro, Portugal; [4] Departamento de Promoção da Saúde e Prevenção das Doenças Não Transmissíveis, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal

Background

Metabolic syndrome (MetS) is a cluster of conditions — increased blood pressure, high blood glucose level, excess body fat around the waist and abnormal cholesterol levels — that occur together, increasing the risk of heart disease, stroke and diabetes. In Portugal, its prevalence is estimated to be 27.5%, constituting a public health problem [1]. As a complex condition, MetS results from a complex interplay between many genetic and environmental factors. Genome-wide association studies (GWAS) have identified various SNPs associated with MetS traits but to date, no loci have been found affecting its entire spectrum. Because pathways and processes implicated in different diseases reveal surprising insights into shared genetic bases underlying apparently unrelated traits, we hypothesize that there is also a common genetic factor involved in the clustering of MetS traits and, as a consequence, of different cardiovascular risk factors.

Objective

The main objective of this study is to identify and characterize genetic factors involved in MetS clustering etiology, using a Principal Component Analysis (PCA) derived continuous MetS score to perform a genetic association study, considering SNPs in candidate genes related to MetS features.

Materials and Methods

Study design and participants

A cross-sectional study developed in the context of the pilot study of the Portuguese Component of the European Health Examination Survey (EHES) project was used [2,3]. Data was collected in 2010 in the population covered by the São Brás de Alportel (Algarve) Health Center, through a detailed questionnaire, physical examination and blood sample collection. Participants were selected using a simple random sampling scheme from the National Health System card number database. All participants signed an informed consent form and study protocol was approved by the Ethics Committee of National Health Institute Doctor Ricardo Jorge and by the National Commission for Data Protection.

SNP selection and Genotyping

37 SNPs were selected based on their involvement in metabolic related phenotypes: glucose/insulin homeostasis (*CDKAL1* rs7754840; *CDKN2A/B* rs10811661; *HHEX* rs1111875; *IGF2BP2* rs4402960; *IL6* rs1800795; *KCNJ11* rs5219; *KCNQ1* rs2237892; *MTNR1B* rs10830963; *PPARG* rs1801282; *SLC30A8* rs13266634; *TCF7L2* rs7903146; *ADCY5* rs11708067 and *KCNQ1* rs231362), body mass index (*GNPDA2* rs10938397; *MTCH2* rs10838738; *NPC1* rs1805081; *PTER* rs10508503; *SH2B1* rs7498665; *FTO* rs9939609; *ADRB3* rs4994; *GABRA2* rs279871; *NPY* rs16147; *TMEM18* rs6548238), cardiovascular system regulation (*ACE* rs4646994; *NOS1AP* rs12143842; *ADRB1* rs1801252; *ADRB2* rs1042714; *ADRB2* rs1042713; *NOS3* rs1799983; *NOS3* rs2070744) and drug/lipid metabolism (*APOE* rs7412; *LDLR* rs2228671; *CYP2C8* rs10509681; *CYP2C9* rs1799853; *CYP2D6* rs16947; *CYP2C19* rs4244285; *TPMT* rs1142345). All SNPs were typed by Sequenom Mass ARRAY platform except 5 (rs1801252, rs10509681, rs16947, rs2070744 and rs464699) that were genotyped by RFLPs analysis.

Statistical analysis

The statistical analysis was performed using *IBM SPSS statistics 20*. P-values < 0.05 were considered to denote statistical significance. The MetS score was calculated by PCA with varimax rotation [4]. Six quantitative MetS Traits (waist circumference, diastolic blood pressure, systolic blood pressure, Glucose, Triglycerides and high density lipoprotein cholesterol plasma levels) were normalized and used to obtain the MetS score, with a higher MetS score indicating a less favorable MetS profile. Validity of the MetS score was tested using the ANOVA for trend analysis.

All SNPs were tested for the Hardy Weinberg Equilibrium using the χ^2 -test. Association between the MetS score and individual SNPs was tested by T-test. Correction for multiple comparisons was performed using the *Bonferroni* method. ANOVA for trend (p<0.05) was used to assess linearity between the MetS score and the number of genetic risk factors and to test additive genetic effects of risk variants in increasing MetS score values. General linear model (GLM) was used to test MetS score differences between subjects with different genotypes after adjusting for confounding variables such as age and gender.

Results

The final study population consisted on 206 participants, 87 (42.2%) men and 119 (57.8%) women. The participants' age ranged from 26 to 91 years, being the mean value 56.43 ± 16.23. In the total sample, from PCA considering the six quantitative MetS risk factors, we were able to explain 63.35% of these six components variance (PC1 and PC2 explained 35.42% and 27.43% of the variance, respectively). The MetS score adequacy and validity is shown in Figure 1. As expected, this score increases progressively with increasing numbers of risk factors (ANOVA test for linear trend, p<0.001).

Regarding the *CYP2C19* rs4244285 SNP, individuals included in the GA+AA genotype group seem to be protected against MetS, displaying a lower MetS score (Mean difference: 0.792; 95%CI: 0.351-1.233; p<0.001) (Table 1). This association remains significant after *Bonferroni* correction for multiple testing (p=0.018).

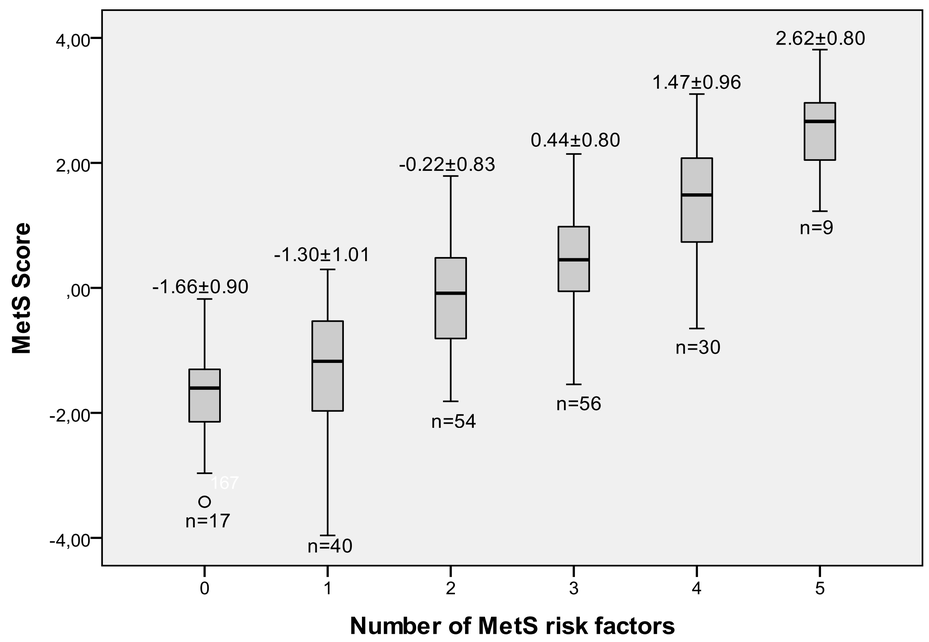

Figure 1- MetS score variation according to the number of risk factors (ANOVA for trend p<0.001).

Using a general linear model, we found that differences on the MetS score between subjects with GG genotype and GA+AA genotype on the *CYP2C19 SNP* remain significant after adjustment for age and gender (Mean difference: 0.768; 95% CI. 0.356-1.180; p=0.011). No association was found between the MetS score and lifestyle risk factors.

Table 1- SNPs significantly associated with MetS score (MetS score are presented as mean ± SD).

Gene	Genotype	n	MetSscore	Mean difference	95% CI	P-value[a]	Corrected P-value[b]
CYP2C19	GG	156	0.192 ± 1.380	0.792	0.351-1.233	0.00049	**0.018**
rs4244285	GA+AA[1]	50	-0.600 ± 1.362				
GABRA2	AA	63	0.350 ± 1.374	0.504	0.087-0.921	0.018	0.670
rs279871	GA+GG[2]	143	-0.154 ± 1.409				
NPY	AA	58	0.342 ± 1.606	0.476	0.048-0.904	0.029	0.999
rs16147	GA+GG[3]	148	-0.134 ± 1.313				
TPMT	AA	192	--0.080 ± 1.375	1.199	0.413-1.984	0.003	0.109
rs1142345	GA	13	1.119 ± 1.601				

[a] T-test was used to compare MetSscore mean values between the two groups.
[b] Corrected P-values were obtained using the *Bonferroni* test to multiple testing correction.
[1]GA+AA group consist in 3 AA and 47 GA individuals; [2]GA+GG group consist in 36 AA and 107 GA individuals;
[3]GA+GG group consist in 37 AA and 111 GA individuals.

An additive genetic effect, age independent, of the *rs279871*, *rs16147* and *rs1142345* SNPs in the *GABRA2*, *NPY* and *TPMT* genes was detected (Figure 2), since MetS score increases with the increasing number of genetic risk factors (ANOVA for trend p<0.001).

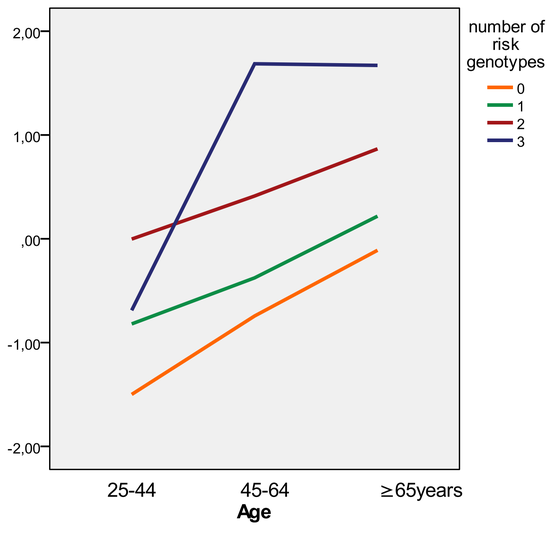

Figure 2- Additive genetic effect representation. Each line represents a different number of risk genotypes, considering sufficient the presence of one allele in each variant. We have considered the 4 significantly associated SNPs previous reported: *CYP2C19* rs4244285, *GABRA2* rs279871, *NPY* rs16147 and *TPMT* rs1142345. No individuals with 4 risk genotypes for the 4 SNPs were identified in this population.

Conclusions

- MetS score explains over 63.0% of the phenotype, supporting the usefulness of a continuous MetS risk score, instead of the dichotomized MetS definition traditionally used in genetic association studies.
- A significant association between *CYP2C19 rs4244285* and the MetS score, which was corrected for multiple testing, was found and it remains significant after adjustment for age and gender.
- Other variants in the *GABRA2*, *NPY* and *TPMT* genes might represent additive genetic factors of modest effect that should be taken in consideration in the etiology of MetS.
- Our results suggest that *CYP2C19* rs4244285 is involved in a common pathway, the deregulation of which, in addition to other specific genetic factors, may lead to the different MetS associated traits.
- Our study represents an integrative approach to identifying genetic risk factors involved in complex disorders' etiology, through continuous scores obtained by PCA.

Bibliography

[1] Fiuza M, Cortez-Dias N, Martins S: **Síndrome metabólica em Portugal: prevalência e implicações no risco cardiovascular-Resultados do Estudo VALSIM**. *Revista portuguesa de cardiologia* 2008, **27**(12):1495-1529.
[2] Kuulasmaa K, Tolonen H, Koponen P, Kilpeläinen K, Avdicová M, Broda G, Calleja N, Dias C, Gösswald A, Kubinova R *et al*: **An overview of the European Health Examination Survey Pilot Joint Action**. *Archives of public health = Archives belges de sante publique* 2012, **70**(1):20.
[3] **Inquérito de Saúde com Exame Físico: Relatório do estudo realizado em São Brás de Alportel.** Lisboa: Instituto Nacional de Saúde Doutor Ricardo Jorge. Departamento de Epidemiologia. Departamento da promoção de saúde e doenças crónicas, 2012. **[http://hdl.handle.net/10400.18/981]**
[4] Wijndaele K, Beunen G, Duvigneaud N, Matton L, Duquet W, Thomis M, Lefevre J, Philippaerts R: **A continuous metabolic syndrome risk score: utility for epidemiological analyses**. *Diabetes Care* 2006, **29**:2329.